

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

**ÉTUDE DES GÈNES DIFFÉRENTIELLEMENT EXPRIMÉS DANS LES
LEUCÉMIES LYMPHOIDES INDUITES PAR LE RÉTROVIRUS MURIN GRAFFI**

**MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN BIOLOGIE**

**PAR
CHARFI CYNDIA**

OCTOBRE 2006

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 -Rév.01-2006). Cette autorisation stipule que «conformément à l'article **11** du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

En premier lieu, je souhaite remercier mon directeur de maîtrise, le Dr. Éric Rassart pour m'avoir accueilli dans son laboratoire. Merci pour la liberté et la confiance qu'il a su m'accorder tout au long de ces deux années. Je pense que je n'aurais pas pu trouver un meilleur environnement pour ma maîtrise.

Je voudrais ensuite remercier Véronique Voisin qui m'accompagne depuis mes premiers pas au laboratoire. Je la remercie pour son aide et pour tout le temps qu'elle m'a consacré.

Un grand merci à toute l'équipe du laboratoire du Dr. Éric Rassart (Sonia Do Carmo, Elsy Édouard, Séverine Landais, Fadila Larguet, Philippe Legault et Louis-Charles Levros) pour leur aide et leur soutien. Je remercie Fadila pour ses conseils toujours pleins de bon sens. Je remercie également Elsy pour son dévouement et sa très grande gentillesse.

Je tiens à remercier tous mes amis, plus spécifiquement, Fatma, Imen et Raoudha. Leur amitié est un cadeau précieux, auquel je tiens énormément.

Je remercie très chaleureusement mon frère pour tout ce qu'il a fait pour moi. Sa douceur et son humour ont favorisé mon acclimatation à ce nouvel environnement et m'ont apporté l'équilibre indispensable à la réussite de ces 2 années de maîtrise. Je remercie également Katia pour sa gentillesse et pour son bon cœur.

Je ne pouvais terminer ces remerciements sans penser à mes très chers parents. Merci pour avoir cru en moi, pour m'avoir inlassablement encouragé et pour avoir toujours été à mes côtés. Merci pour avoir réalisé mon rêve d'enfant en me donnant l'opportunité de poursuivre mes études au Canada. Ces quelques lignes ne pourront certainement pas refléter à quel point je leur suis reconnaissante.

TABLE DES MATIÈRES

LISTE DES FIGURES	viii
LISTE DES TABLEAUX	x
LISTE DES ANNEXES	xii
LISTE DES ABBRÉVIATIONS	xiii
RÉSUMÉ	xvii
INTRODUCTION	1
CHAPITRE I	
ÉTAT DES CONNAISSANCES	3
1.1 Le système hématopoïétique	3
1.1.1 Introduction.....	3
1.1.2 Les cellules souches hématopoïétiques.....	3
1.1.3 L'hématopoïèse.....	4
1.1.3.1 Définition de l'hématopoïèse et facteurs de régulation.....	4
1.1.3.2 Les précurseurs des cellules sanguines de la moelle osseuse	6
1.1.3.2.1 Les précurseurs de la lignée myéloïde : CFU-GEMM	6
1.1.3.2.2 Les précurseurs de la lignée lymphoïde : CFU-L	7
1.1.4 Lymphocytes T et lymphocytes B	8
1.1.4.1 Les lymphocytes T	8
1.1.4.2 Les lymphocytes B	9
1.1.5 Formule leucocytaire de l'adulte normal	10
1.1.6 Les organes hématopoïétiques	11
1.1.6.1 La moelle osseuse	11

1.1.6.2	Les ganglions lymphatiques.....	11
1.1.6.3	La rate	12
1.1.6.4	Le thymus.....	12
1.2	Cancer et leucémies.....	12
1.2.1	Le cancer	12
1.2.2	Les leucémies	13
1.2.2.1	Les différents types de leucémies chez humain	14
1.2.2.2	Différents types de leucémies chez la souris.....	15
1.3	Les rétrovirus	16
1.3.1	Classification et structure des rétrovirus (<i>Retroviridae</i>)	16
1.3.1.1	La famille des Oncovirinae (Oncovirus)	16
1.3.1.2	Les <i>Spumavirinae</i> (Spumavirus)	17
1.3.1.3	Les <i>Lentivirinae</i> (Lentivirus).....	17
1.3.2	Structure des rétrovirus.....	17
1.3.3	Cycle de réplication des rétrovirus	19
1.3.4	Oncogènes	21
1.3.5	Activation des oncogènes.....	22
1.3.6	Oncogènes et leucémies humaines	23
1.3.7	Le rétrovirus murin Graffi.....	24
1.4	Les micropuces à ADN : Méthode d'analyse de l'expression génique des cellules leucémiques et non leucémiques	25
1.5	Hypothèse de travail.....	27

CHAPITRE II

MATÉRIELS ET MÉTHODES 29

2.1 Retrovirus murin Graffi..... 29

2.2 Les organes tumoraux 29

2.3 Les cellules leucémiques..... 30

2.4 La cytométrie de flux (FACScan).....30

2.5 Le tri des cellules par la technique des billes magnétiques..... 33

2.6 Analyses en biologie moléculaire 34

2.6.1 Etude de l'expression génique de quelques gènes par Northern Blot 34

2.6.2 Hybridation de la sonde radioactive pour le Northern Blot..... 34

2.7 Les micropuces : Méthode d'analyse de l'expression génique des cellules
de souris leucémiques et non leucémiques..... 35

2.8 RMA et sélection des gènes 39

2.9 Sélection de quelques gènes pour confirmer le résultat obtenu par
les micropuces à ADN..... 40

2.10 La Réverse Transcriptase-PCR (RT-PCR)..... 44

2.10.1 Quantification du niveau d'expression des gènes GATA-1, GATA-2
et GATA-3 par Reverse Transcriptase-PCR semi-quantitative 45

2.10.2 Confirmation du résultat des micropuces par
Reverse Transcriptase-PCR..... 46

CHAPITRE III

RÉSULTATS 48

3.1 Caractérisation des leucémies 48

3.1.1 Analyse des signes cliniques chez la souris et de la morphologie
des organes..... 48

3.1.2	Analyse des cellules leucémiques par cytométrie de flux	49
3.1.3	Analyse moléculaire.....	56
3.1.3.1	Etude de l'expression du gène Fog1 par Northern blot.....	56
3.1.3.2	Etude du niveau d'expression des gènes GATA-1, GATA-2 et GATA-3 par RT-PCR semi-quantitative.....	58
3.1.3.2.1	Mise au point de la technique de la RT-PCR semi-quantitative pour GATA-1, GATA-2 et GATA-3	59
3.1.3.2.2	Etude du niveau d'expression des gènes GATA-1, GATA-2 et GATA-3 au niveau de différents types de leucémies par RT-PCR semi-quantitative	61
3.2	Analyse des différentes leucémies induites par le rétrovirus Graffi à l'aide des micropuces à ADN.....	66
3.2.1	Tri des cellules leucémiques selon leurs caractéristiques phénotypiques	66
3.2.2	Micropuces à ADN	68
3.2.2.1	Première sélection des gènes d'intérêt.....	68
3.2.2.2	Méthode de sélection par «Hierarchical Clustering»	69
3.2.2.3	Analyse des gènes sélectionnés	71
3.2.2.4	Etude de la variation intra-leucémique de l'expression de certains gènes	72
3.2.2.4.1	Analyse par RMA	72
3.2.2.4.2	Analyse par «Scatter plots MAS 5.0» (diagrammes de dispersion).....	73
3.3	Validation des résultats obtenus suite à l'analyse par les micropuces par RT-PCR.....	76
3.3.1	Choix des gènes à analyser.....	76

3.3.2 Validation du résultat des micropuces à ADN par RT-PCR	79
--	----

CHAPITRE IV

DISCUSSION	91
-------------------------	-----------

CHAPITRE V

CONCLUSIONS ET PERSPECTIVES FUTURES	111
--	------------

ANNEXES	115
----------------------	------------

BIBLIOGRAPHIE	124
----------------------------	------------

LISTE DES FIGURES

Figure	Page
1.1 Schéma de l'hématopoïèse	5
1.2 Visualisation de la structure du rétrovirus	18
1.3 Représentation schématique du rétrovirus	18
1.4 Organisation de l'ARN proviral d'un rétrovirus	19
1.5 Mécanisme de répllication des rétrovirus	21
1.6 Principe de la méthode des micropuces à ADN « Microarrays »	27
2.1 Schématisation des différentes étapes nécessaires pour la préparation et l'hybridation des fragments d'ARN sur les micropuces à ADN	37
2.2 Localisation de la séquence Affymétrie au niveau du transcrit du gène Arntl2	42
2.3 Localisation de la séquence Affymétrie au niveau du transcrit du gène Gpm6a	43
2.4 Localisation de la séquence Affymétrie au niveau du transcrit du gène Neil3	43
2.5 Localisation de la séquence Affymétrie au niveau du transcrit du gène Dir2	44
3.1 Immunophénotypage des cellules leucémiques et normales par cytométrie de flux	55
3.2 Analyse de l'expression du gène Fog1 par Northern blot pour différents types de leucémies	58
3.3 Optimisation de la réaction de PCR pour les gènes GATA-1, GATA-2 et GATA-3	60

3.4	Etude du niveau d'expression du gène GATA-1 au niveau de différents types de leucémies	63
3.5	Etude du niveau d'expression du gène GATA-2 au niveau de différents types de leucémies	64
3.6	Etude du niveau d'expression du gène GATA-3 au niveau de différents types de leucémies	65
3.7	Exemple d'un tri d'une leucémie CD4+ CD8- par la technique des billes magnétiques	66
3.8	Algorithme du « Hierarchical Clustering » des leucémies T et B	70
3.9	Comparaisons intra et inter-leucémique à l'aide du diagramme de disparition	74
3.10	Etudes de l'expression de quelques gènes par RT-PCR au niveau des leucémies non triées	81
3.11	Etudes de l'expression des 7 gènes sélectionnés au niveau de leucémies triées.....	87
4.1	Niveau d'expression de GATA-1 et de GATA-2 en fonction de l'érythropoïèse	94

LISTE DES TABLEAUX

Tableau	Page
1.1 Proportions de globules blancs chez l'homme (pour un adulte normal).....	10
1.2 Classification des leucémies chez la souris	15
2.1 Différents anticorps utilisés pour l'immunophénotypage des leucémies par cytométrie de flux et pour le tri des cellules par les billes magnétiques	32
2.2 Caractéristiques phénotypiques des échantillons analysés par les micropuces à ADN	38
2.3 Gènes sélectionnés pour valider le résultat des micropuces à ADN.....	41
2.4 Différentes températures et différents cycles utilisés pour chaque paire d'amorce	46
2.5 Séquence des paires d'amorces utilisées pour chacun des 7 gènes testés par RT-PCR	47
3.1 Liste des souris étudiées et des caractéristiques des tumeurs induites par le virus Graffi.....	51
3.2 Ensemble des leucémies triées par la technique des billes magnétiques	67
3.3 Tableau récapitulatif du nombre de gènes sélectionnés par la méthode RMA.....	69
3.4 Niveau d'expression des gènes connus associés à des leucémies T ou B	70
3.5 Nombre de gènes exprimés en fonction du type de leucémie.....	73

3.6	Niveau d'expression des 7 gènes sélectionnés dans les différentes leucémies d'après les micropuces à ADN	76
3.7	Leucémies non triées	79
3.8	Ensemble des leucémies triées.....	84

LISTE DES ANNEXES

Annexe		Page
Annexe A	Pathogénèse du rétrovirus murin Graffi	115
Annexe B	Caractéristiques des leucémies choisies pour tester le niveau d'expression de GATA-1, GATA-2 et de GATA-3.....	116
Annexe C	Figure récapitulant le nombre de gènes qui a été trouvé dans chaque groupe de leucémies	118
Annexe D	Gènes spécifiques aux leucémies B	119
Annexe E	Gènes spécifiques aux leucémies T.....	120
Annexe F	Gènes communs aux leucémies T et B.....	121
Annexe G	Fonctions biologiques des 7 gènes sélectionnés pour confirmer le résultat des micropuces à ADN	122
Annexe H	Résultats obtenus par MAS5 suite à l'analyse de l'expression des 7 gènes	123

LISTE DES ABBRÉVIATIONS

°C	Degré Celsius
µl	Microlitres
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
Ag	Antigène
ALL	Leucémies aïgues lymphoblastiques
AML	Leucémies aïgues myéloblastiques
Arhgap8	« Rho GTPase activating protein 8 »
ARN	Acide ribonucléique
ARNm	Acide ribonucléique messenger
Arntl2	« Ayl hydrocarbon receptor nuclear translocator-like 2 »
ATL	« Adult T-cell leukemia »
B	Leucémie lymphoïde de type B
B-ALL	Leucémies aïgues lymphoblastiques de type B
B-CLL	Leucémies chronique de type B
Bfsp2	« Beaded filament structural protein 2 »
C	Cerveau
C3d	« Complement Component 3d »
CA	Capside
Cabp4	« Calcium binding protein 4 »
CB	Contrôle pour les leucémies B
Ccr1	« Chemokine (C-C motif) receptor 1 »
CD	« Cluster of differentiation »
Celsr1	« Cadherin EGF LAG seven-pass G-type receptor 1 »
CFU	« Colony forming unit »
CLL	Leucémies chroniques lymphoïdes
CML	Leucémies chroniques myéloïdes
Crisp3	« Cysteine-rich secretory protein 3 »
CT	Contrôle pour les leucémies T
Dirc2	« Disrupted in renal carcinoma 2 (human) »
Dntp	Deoxyribonucleotide triphosphate
E	Leucémie érythroïde
EGF	« Epidermal Growth Factor »
Elovl6	« ELOVL family member 6, elongation of long chain fatty acids (yeast) »
Enpep	« Glutamyl aminopeptidase (aminopeptidase A) »
Env	Gène de l'enveloppe
ERM	« Ets-related molecule PEA3-like »
EST	« Expressed sequence tag »
Etv5	« Ets variant gene 5 »
F	Foie
FCS	Sérum de veau foetal
FGF	«Fibroblast Growth Factors»

FITC	Fluorescéine
Fmn2	« Formin 2 »
Fog1	« Friend of GATA-1 »
FSC	« Forward scatter »
G	Ganglions
G.B	Globules blancs
G.R	Globules rouges
Gag	Gène de la capside
GATA-1	« GATA binding protein 1 »
GATA-2	« GATA binding protein 1 »
GATA-3	« GATA binding protein 1 »
Gpm6a	« Glycoprotein m6a »
Gpm6b	« Glycoprotein m6b »
HCL	« Hairy cell leukemia »
Hectd2	« HECT domain containing 2 »
Hist1h2bc	Histone 1, H2bc
Hmgcll1	« 3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase-like 1 »
HTLV-1	« human T-cell leukemia virus type 1 »
HTLV-2	« human T-cell leukemia virus type 2 »
IgM	Immunoglobulines M
IP	Iodure de propidium
JH	Chaînes lourdes d'immunoglobulines
kD	Kilo Dalton
KHCO3	Bicarbonate de potassium
L.B	Lymphocytes B
L.T	Lymphocytes T
Lef1	« Lymphoid enhancer binding factor 1 »
Log	Logarithme
LTR	« Long Terminal Repeat »
M	Leucémie myéloïde
Ma	Matrice virale
Map4k5	« Mitogen-activated protein kinase kinase kinase kinase 5 »
MCL	« Mantle cell lymphoma »
Mettl1	« Methyltransferase-like 1 »
Mg	Milligramme
Min	Minute
Mk	Leucémie mégacaryocytaire
Mki67	« Antigen identified by monoclonal antibody Ki 67 »
mL	Millilitre
MM	« MisMatch »
MuLV	« Murine Leukemia Virus »
MZL	« Marginal zone B-cell lymphoma »
Na4EDTA	Ethylenediamine Tetraacetic Acid tétrasodique
NaCl	Chlorure de sodium
NaH ₂ PO ₄	Phosphate de sodium monobasique
NaOAc	Acétate de sodium

NaOH	Hydroxyde de sodium
Neil3	« Neil-like 3 (<i>E. coli</i>) »
NH ₄ Cl	Chlorure d'ammonium
NK	« Natural killer »
Nln	« Neurolysin (metallopeptidase M3 family) »
Notch-1	« Notch Gene Homolog 1 »
Notch3	« Notch gene homolog 3 »
Nrxn1	« Neurexin I »
NU	Nucléocapside
Olfml2b	« Olfactomedin-like 2B »
OMS	Organisation Mondiale de la Santé
Pb	Paire de bases
PBS	Tampon phosphate (PBS «Phosphate Buffered Saline»)
PCP	« Planar cell polarity »
PCR	« Polymerase chain reaction »
PE	« Phycoérythrine »
Pkib	« Protein kinase inhibitor beta »
Pla2g4a	« hypoxia involved Pla2g4a gene »
PLP	« Myelin proteolipid protein »
PM	« Perfect Match »
Pol	Gène de transcriptase inverse (RT), protéase et intégrase
PPT	« Polypurine Tract »
R	Rate
Rag2	« Recombination activating gene 2 »
RMA	« Robust multichip average »
RT-PCR	« Reverse Transcriptase- Polymerase Chain Reaction »
S	Sang
SDS	« Sodium dodecyl sulfate »
Serpinf1	« Serine (or cysteine) peptidase inhibitor, clade F, member 1 »
SLL	« Small lymphocytic lymphoma »
SMS-SB	« Human pre-B cell-like acute lymphocytic leukemia cell line »
SSC	« salin citrate de sodium »
SSC	« side scatter »
SSPE	« Saline Sodium Phosphate EDTA »
SU	Glycoprotéine de la surface de l'enveloppe virale
T	Leucémie lymphoïde de type T
T-ALL	Leucémie lymphoblastiques aigue des cellules T
TCR	« T-cell receptor »
Tcra	« T-cell receptor alpha chain »
TdT	« Terminal désoxynucléotidyl transférase »
Thy	Thymus
Tm	« Melting temperature »
TM	Glycoprotéines transmembranaires
Tmc7	« Transmembrane channel-like gene family 7 »
U3	Séquence unique dans le LTR du génome viral
U5	Séquence unique dans le LTR du génome viral

UTR	« Untranslated region »
VLCFA	« Very long chain fatty acids »
µg	Microgramme

RESUMÉ

Les leucémies lymphoïdes de type T et B sont des maladies malignes typiques des cellules sanguines. Afin d'approfondir nos connaissances concernant ces dernières, nous avons utilisé le rétrovirus murin Graffi. Pendant longtemps, on pensait que ce rétrovirus ne provoquait que l'apparition des leucémies myéloïdes pour constater, plus tard avec des outils moléculaires et plus perfectionnés, qu'il était capable d'induire différents types de leucémies (lymphoïdes et non lymphoïdes). Pour mieux caractériser les leucémies lymphoïdes et dans le but de trouver des gènes potentiellement impliqués dans ce type de leucémie ou dans le processus de l'hématopoïèse, la technique des microarrays a été retenue. Cette technique permet de mesurer en une seule expérience le niveau d'expression de plusieurs milliers de gènes. Les analyses ont porté sur 8 échantillons de souris constitués de 3 sous types de leucémies T ($CD4^+/CD8^+$, $CD4^+/CD8^-$, $CD4^-/CD8^+$), de 3 sous-types de leucémies B ($CD45R^{\text{low}}/CD19^+$, $CD45R^+/CD19^+$, $CD45R^+/CD19^+/SCA1^+$) et d'un contrôle constitué de lymphocytes T ($CD4^+/CD8^+$) et B ($CD45R^+/CD19^+$) exprimant les marqueurs de surface exprimés à la surface de chacun des types de leucémies. Différentes approches d'analyse des résultats obtenus par les microarrays ont été appliquées de façon à distinguer ou à regrouper les gènes dont l'expression est altérée dans les différents types de leucémies analysées. Ceci nous a permis de sélectionner 56 gènes qui étaient soit spécifiques aux leucémies T, soit spécifiques aux leucémies B, soit communs à ces deux types. Parmi ces gènes se trouvent des gènes qui sont connus dans d'autres types de cancers mais qui n'ont pas encore été impliqués dans la leucémie. D'autres ne sont connus que dans les leucémies non lymphoïdes ou encore n'ont jamais été impliqués ni dans la tumorigénèse ni dans la leucémogénèse mais qui ont des fonctions physiologiques nécessaires au bon fonctionnement de l'organisme. Pour certains autres gènes, aucune information n'était disponible. L'application des microarrays a donc permis de créer une liste constituée des gènes potentiellement impliqués dans le développement de leucémies et qui pourraient servir de marqueurs diagnostiques.

Mots-clés : Rétrovirus, rétrovirus murin Graffi, leucémies lymphoïdes, leucémies non lymphoïdes, cytométrie de flux, tri cellulaire, micropuces à ADN, gènes différentiellement exprimés, oncogènes.

INTRODUCTION

La leucémie est une maladie qui s'accompagne d'une dérégulation au niveau du processus de l'hématopoïèse, mécanisme de production des cellules sanguines (globules blancs, globules rouges et plaquettes). En fait elle est due à la présence excessive d'une famille anormale et maligne de cellules de la moelle osseuse, où sont fabriquées les cellules du sang, qui, par la suite, vont infiltrer plusieurs autres organes dont le foie et la rate.

Pour comprendre ce phénomène chez l'humain, il a fallu le reproduire expérimentalement chez un modèle animal ayant les mêmes caractéristiques que l'homme. Vu l'étendue de la ressemblance génétique entre l'homme et la souris, cette dernière a été considérée comme le mammifère le plus accessible et le modèle idéal à utiliser à cet effet permettant ainsi l'étude des conditions physiologiques et des pathologies moléculaires du cancer humain (Bernardi et al. , 2002).

Le rétrovirus murin Graffi nous a paru être un outil précieux pour étudier la leucémie chez la souris. Le choix de cet outil biologique a été basé sur le fait que le rétrovirus murin Graffi a la capacité d'induire la leucémie chez la souris en une période de temps relativement courte d'environ 2 à 4 mois.

Pendant longtemps, on pensait que le rétrovirus Graffi ne provoquait l'apparition que de leucémies de type myéloïde faisant partie des leucémies non lymphoïdes. Plus tard, il s'est avéré que ce rétrovirus était capable d'induire différents types de leucémies d'une façon aléatoire (Voisin et al. , 2006).

Ce qui est encore plus intéressant, c'est qu'il a été démontré que ce rétrovirus est capable d'induire des leucémies mégakaryocytaires qui sont très rares et mal connues. Ce modèle expérimental créé par l'infection de souris nouveau-nées par le rétrovirus murin Graffi, nous offre la possibilité d'approfondir nos connaissances concernant les mécanismes d'induction de la leucémie ou encore du cancer.

Dans le cadre de cette étude, nous avons pensé qu'il était utile de s'intéresser à un tel sujet focalisant notre analyse sur ces deux types de leucémies afin de déterminer les changements génomiques responsables de la conversion des cellules normales en cellules malignes. Dans le but de découvrir de nouveaux gènes impliqués dans la transformation par le rétrovirus Graffi ou des gènes impliqués dans le processus de l'hématopoïèse et à l'aide de la puissante technique des micropuces à ADN (microarrays), nous avons pu caractériser les leucémies T et B induites par le rétrovirus murin Graffi, en analysant la totalité de leur expression génique. Cette technique nous a donc permis de découvrir un ensemble de gènes qui jusqu'à présent n'ont pas encore été associés aux leucémies.

CHAPITRE I

ETAT DES CONNAISSANCES

1.1 Le système hématopoïétique

1.1.1 Introduction

Les cellules sanguines sont représentées par deux types de cellules : les érythrocytes et les leucocytes. Les érythrocytes ou encore les globules rouges (G.R) sont responsables du transport de l'oxygène et du CO₂. Par contre, les leucocytes ou globules blancs (G.B) jouent un rôle primordial dans la défense du système immunitaire contre tout type d'agression. A cause de leur durée de vie limitée et dans le but de maintenir leur concentration sanguine constante, ces cellules doivent être renouvelées quotidiennement au niveau des organes hématopoïétiques. Ce sont les cellules souches hématopoïétiques qui sont à la base de cette production continue et hautement contrôlée.

1.1.2 Les cellules souches hématopoïétiques

La cellule souche hématopoïétique est à l'origine des différentes populations cellulaires (Allen et Dexter, 1995). Ce n'est qu'en 1995 que cette cellule a été purifiée chez la souris.

Chez l'homme, tout au début de la gestation, les premières cellules souches hématopoïétiques peuvent être mises en évidence dans le foie, dans la rate et dans la moelle osseuse (Choquet, 2002). Après la naissance, ces cellules colonisent la moelle osseuse, site exclusif de l'hématopoïèse. Cependant, chez la souris adulte, la rate reste un organe hématopoïétique contrairement à l'homme (Fredrickson et Harris, 2000). Récemment les chercheurs ont découvert la présence de rares cellules souches au niveau du sang périphérique et du sang du cordon ombilical (Chabannon et Mannouni, 1995).

La cellule souche hématopoïétique possède deux propriétés essentielles non rencontrées au niveau des stades ultérieurs de différenciation.

- La première est sa capacité d'auto-renouvellement lui permettant de se maintenir en nombre constant.
- La deuxième est sa capacité de différenciation en tout type de cellules sanguines périphériques matures assurant le renouvellement des cellules mortes et l'accroissement du renouvellement en cas de besoin, d'où sa qualification de cellule pluripotente. Cette pluripotentialité détermine sa capacité de différenciation en au moins 9 lignées cellulaires : les érythrocytes, les neutrophiles, les éosinophiles, les basophiles, les monocytes, les ostéoclastes, les mégakaryocytes, les lymphocytes T (L.T) et les lymphocytes B (L.B) (Allen et Dexter, 1995).

1.1.3 L'hématopoïèse

1.1.3.1 Définition de l'hématopoïèse et facteurs de régulation

L'hématopoïèse est la fonction de production et de renouvellement continu des éléments figurés du sang (les G.B et les G.R ainsi que les plaquettes). C'est un processus coordonné du point de vue prolifération, différenciation et apoptose des cellules hématopoïétiques.

Il existe plusieurs facteurs de régulation qui interviennent dans le bon déroulement des différentes étapes de l'hématopoïèse (Allen et Dexter, 1995, Dittel t al. , 1993). Nous pouvons citer les facteurs diffusibles ou cytokines (comme l'interleukine 1, 3 et 6 ou les «colony-stimulating factors» (G-CSF, M-CSF et GM-CSF)) qui agissent sur les récepteurs des cellules cibles. Un schéma récapitulatif de l'hématopoïèse est présenté ci-dessous par la figure 1.1.

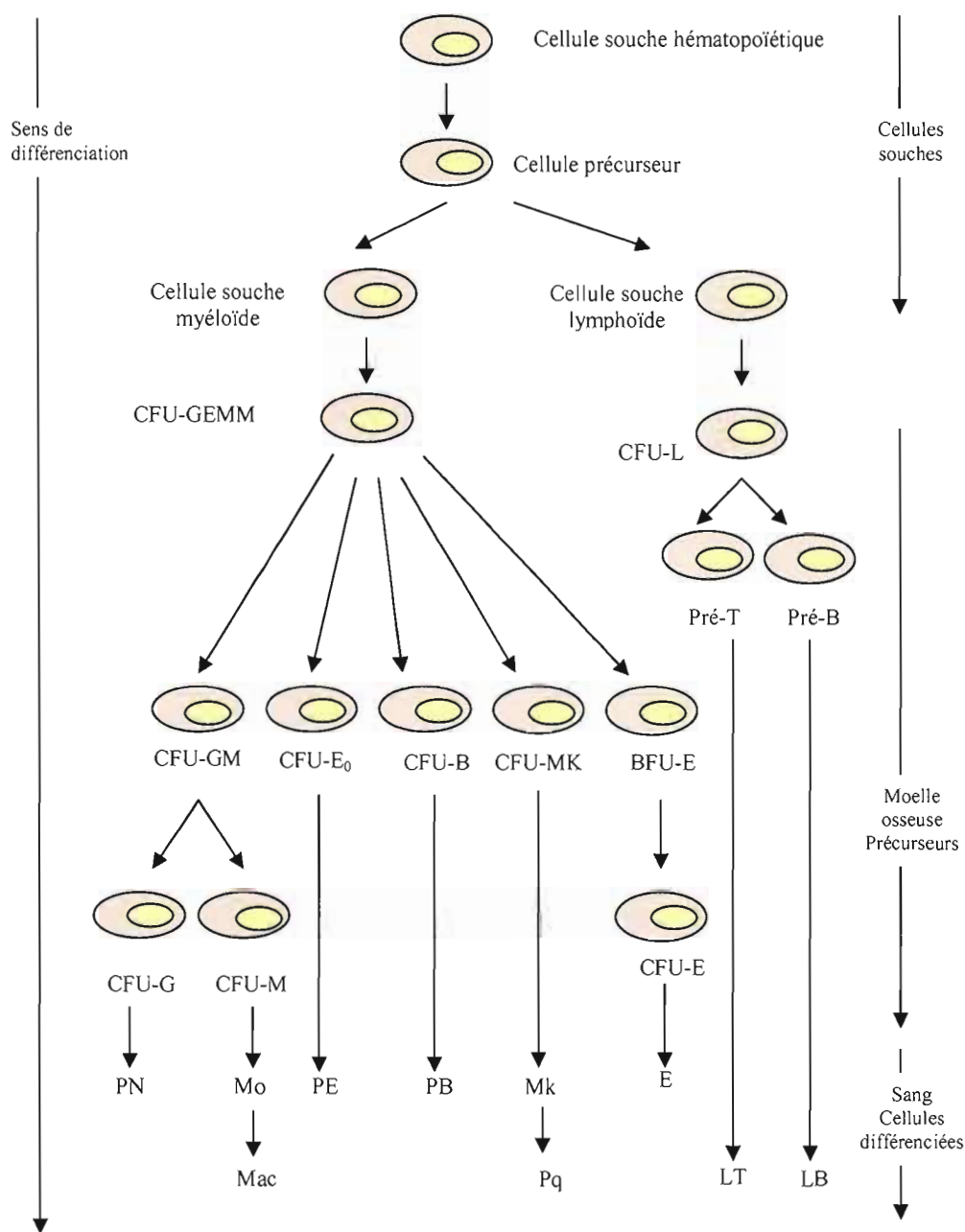


Figure 1.1 Schéma de l'hématopoïèse. Les cellules souches hématopoïétiques donnent naissance à plusieurs progéniteurs (CF-GEMM et CFU-L) qui peuvent se différencier en plusieurs lignées hématopoïétiques. Les CF-GEMM sont à l'origine de la production des polynucléaires neutrophiles (PN), des monocytes (Mo) puis des macrophages (Mac), des polynucléaires éosinophiles (PE) et basophiles (PB), des mégakaryocytes (Mk) puis des plaquettes (Pq) et des érythrocytes (E). Les CFU-L sont à l'origine de la production des lymphocytes T (LT) et B (LB).

Chez la souris et chez l'homme, il a été montré que c'est probablement la perte ou l'acquisition de récepteurs spécifiques pour les facteurs de croissance exogènes qui intervient dans le devenir des cellules souches hématopoïétiques.

1.1.3.2 Les précurseurs des cellules sanguines de la moelle osseuse

Comme la moelle osseuse est le principal lieu de l'hématopoïèse, les divers types de cellules sanguines s'y trouvant peuvent être présentes à différents niveaux de maturation. Une étape importante séparant les cellules souches des cellules différenciées est représentée par l'existence des cellules en voie de différenciation, appelées précurseurs et notées CFU («colony forming unit») en raison de leur aptitude à former des colonies. D'après la figure 1.1, nous pouvons noter l'existence de différents précurseurs l'hématopoïétiques :

1.1.3.2.1 Les précurseurs de la ligné myéloïde : CFU-GEMM (GEMM = Granuleuse, Erythrocytaire, Macrophagique et Mégakaryocytaire) sont des progéniteurs exprimant plusieurs potentialités capables de donner naissance aux :

- **Précurseurs de la lignée granuleuse :CFU-GM** qui, à leur tour, génèrent des précurseurs plus matures qui sont :
 - **les CFU-M** qui, tout au long de leur différenciation, se transforment en monoblastes puis en promonocytes pour donner naissance aux colonies monocytiques. Dès leur passage au niveau tissulaire, les monocytes circulant dans le sang se transforment en macrophages. Ces monocytes et ces macrophages possèdent deux rôles : la cytotoxicité et la présentation des antigènes (Choquet, 2002).
 - **les CFU-G** qui, au cours de leur différenciation se transforment en myéloblastes, en promyélocytes, en myélocytes, en métamyélocytes pour finalement donner naissance aux neutrophiles polynucléaires (Choquet, 2002). Les neutrophiles représentent la lignée majoritaire dans l'organisme

traduisant leur rôle majeur dans la destruction de tout agent pathogène qui a pénétré dans l'organisme (Choquet, 2002).

- **Les CFU-E₀** qui, pour donner naissance aux polynucléaires éosinophiles, se différencient en promyélocyte éosinophile puis en myélocyte éosinophile (Choquet, 2002). L'éosinophile possède plusieurs fonctions comme la destruction des pathogènes par phagocytose ou par libération de toxines, l'inflammation, l'hypersensibilité.
- **Les CFU-B** se différencient en myéloblastes basophiles puis en promyélocyte et finalement en myélocyte basophile avant de donner naissance aux polynucléaires basophiles (Choquet, 2002). Ces basophiles sont très rares. Leur principal rôle est l'hypersensibilité et peuvent aussi déclencher l'anaphylaxie par relargage d'histamine.
- **Précurseurs de la lignée érythroïde : BFU-E (Burst Forming Unit Erythroid)** qui sont moins matures que **les précurseurs CFU-E** et se différencient en proérythroblastes, puis en érythroblastes basophiles, en érythroblastes polychromatophiles et finalement en réticulocytes avant de donner naissance à des colonies érythroïdes (hématies) (Choquet, 2002).
- **Les précurseurs de la lignée mégakaryocytaire : CFU-MK** qui, pour donner naissance à des colonies mégakaryocytaires, se différencient d'abord en mégakaryoblastes, puis en mégakaryocytes basophiles, puis en mégakaryocytes granuleux et finalement en plaquettes (Choquet, 2002). Ces plaquettes constituent la première ligne de défense contre les hémorragies.

1.1.3.2.2 Les précurseurs de la lignée lymphoïde : CFU-L donnent naissance à deux types de cellules : les L.T. et les L.B. Les précurseurs des L.T. quittent la moelle osseuse pour se différencier au niveau du thymus où l'on obtient les thymocytes. Ces L.T. se différencient au niveau du thymus en cellules auxiliaires (qui sont CD4⁺) et en cellules tueuses (qui sont CD8⁺). Les L.B. se différencient au niveau de la moelle osseuse en pro-B

puis en pré-B et enfin en B matures exprimant les immunoglobulines M (IgM) de surface (Choquet, 2002).

Mis en évidence en 1962 par Till et Macculloch, tous ces précurseurs subissent une vingtaine de divisions avant de passer d'une cellule souche qui leur a donné naissance à des cellules complètement différenciées et par une différenciation progressive au cours de laquelle les précurseurs engagés perdent leur capacité d'auto- renouvellement formant, ainsi, des cellules précurseurs mono, bi ou tripotentes. Ces cellules seront à l'origine des cellules hématopoïétiques matures après quelques étapes de différenciation (Allen et Dexter, 1995).

Remarque : Certains précurseurs présents dans le foie fœtal ainsi que dans la moelle osseuse de l'adulte ont été montrés capables de produire une seule colonie de lymphocytes B et de macrophages. Ceci peut refléter la relation existant entre ces lignées ainsi que la plasticité des cellules souches pluripotentes (Palmer et Morrow, 2004; Friedman, 2002).

1.1.4 Lymphocytes T et lymphocytes B

Les lymphocytes T et B diffèrent par leur site de différenciation, par la nature de leur récepteur membranaire pour l'antigène (Ag), par le mode de reconnaissance, par leurs marqueurs membranaires et par leurs fonctions. A la différence de la plupart des autres cellules de l'organisme, les lymphocytes ne peuvent atteindre le stade terminal de leur différenciation qu'à la faveur d'une stimulation par un Ag.

1.1.4.1 Les lymphocytes T (T pour thymus) représentent les cellules de l'immunité à médiation cellulaire (cytotoxicité et hypersensibilité retardée). Les lymphocytes T se différencient dans le thymus à partir de cellules souches lymphoïdes qui proviennent de la moelle osseuse (ou du foie fœtal pendant les premiers mois de gestation). Le réarrangement des gènes des TCR constitue l'évènement majeur au cours de la différenciation intrathymique des cellules T. La filiation intrathymique des cellules de la lignée T peut être subdivisée en 3 grandes étapes reliées à l'expression ou non des marqueurs CD4 et CD8 :

- ♣ **Les précurseurs T encore appelés prothymocytes ou thymocytes précoces** : qui sont doubles négatifs $CD4^-/CD8^-$. Ils sont TdT^+ (Terminal Désoxynucléotidyl Transférase), $Rag1^+$ et $Rag2^+$. Ils expriment aussi la chaîne alpha du récepteur à l'IL2 et les molécules CD2, CD5 et CD7. Ces cellules ont entamé le réarrangement de leurs gènes du TCR.
- ♣ **Les thymocytes communs ou immatures** : qui sont doubles positives $CD4^+/CD8^+$. Elles sont $CD25^-$ mais $CD1^+$, $CD2^+$, $CD5^+$, $CD7^+$.
- ♣ **Les thymocytes matures** : qui sont simples positives $CD4^+/CD8^-$ ou $CD4^- CD8^+$. Ces cellules sont TdT^- , $Rag1^-$ et $Rag2^-$. Elles expriment le complexe TCR/CD3. Elles sont $CD1^+$, $CD2^+$, $CD5^+$, $CD7^+$.

1.1.4.2 Les lymphocytes B (B pour bourse de Fabricius) représentent les cellules de l'immunité humorale médiée par les anticorps. Chez l'homme, comme chez la souris, la production des lymphocytes B se situe au niveau de la moelle osseuse. En fait, dès la gestation, cette moelle osseuse se charge de ce rôle qui était octroyé au foie. Le réarrangement des gènes des Ig représente l'événement majeur au cours de l'ontogenèse des lymphocytes B (les gènes des Ig et du TCR des cellules souches lymphoïdes ont une configuration germinale qui est non réarrangée). Au cours de sa différenciation, la cellule souche lymphoïde engagée dans la lignée B passe successivement par plusieurs étapes. Ces étapes sont caractérisées par l'expression d'un certain nombre de marqueurs membranaires :

- ♣ **pro-B** : à la fin de ce stade, le réarrangement des gènes des Ig est entamé au niveau des gènes des chaînes lourdes par un réarrangement $D \rightarrow JH$. Ces cellules expriment les molécules HLA de classe II. Elles sont aussi $CD19^+$, $CD10^+$, TdT^+ , $Rag1^+$, $Rag2^+$, $CD40^+$, $CD45R^+$, $CD21^+$ et $CD22^+$, $CD38^+$.
- ♣ **pré-B** : à ce stade, le réarrangement des gènes des chaînes lourdes est achevé ($V \rightarrow DJH$). Ces cellules sont $CD10^-$, TdT^- , $Rag1^-$, $Rag2^-$, $V \text{ pré-B}^-$ et $\lambda 5^-$, $CD38^+$. C'est à partir de ce stade que commence l'apparition des gènes Ig α et Ig β .
- ♣ **B immatures** : ce stade est caractérisé par l'expression du récepteur spécifique pour l'Ag du lymphocyte B ou BCR.

- ♣ **B matures au repos (vierge, naïf)** : ils expriment l'IgD membranaire et sont CD35⁺ et CD32⁺.
- ♣ **B actifs** : ils sont caractérisés par la perte de l'IgD membranaire. Ils sont CD23⁻, CD80⁺ et CD10⁺.
- ♣ **B mémoires** : ils sont IgM⁻, CD23⁻, CD80⁻, CD10⁻, iGg⁺, IgA⁺.
- ♣ **Plasmocytes** : ils sont caractérisés par la perte de tous les marqueurs membranaires sus-cités y compris les Ig de surface et les Ag HLA classe II. Ils sont CD38⁺.

1.1.5 Formule leucocytaire de l'adulte normal

Le tableau suivant représente les proportions respectives des globules blancs chez l'homme (pour un adulte normal) (tiré de Lord Dubé et l'Italien, 1983) :

Tableau 1.1 Proportions de globules blancs chez l'homme (pour un adulte normal)

Type de leucocytes	%	Nombre absolus (cellules x 10 ⁹ /L)
Granulocytes neutrophiles	40 – 75	2 - 7,2
Granulocytes éosinophiles	1 – 2	0,05 – 0,2
Granulocytes basophiles	0 – 1	0,02 – 0,06
Lymphocytes	25 – 33	1,5 – 3,5
Monocytes	4 – 8	0,2 – 0,8

Les données présentées dans ce tableau montrent que les éosinophiles, mais surtout les basophiles, sont très rares alors que les neutrophiles représentent une fraction importante des cellules leucocytaires.

1.1.6 Les organes hématopoïétiques

La moelle osseuse, les ganglions, la rate et le thymus représentent les organes hématopoïétiques.

1.1.6.1 La moelle osseuse

La moelle osseuse est le lieu de production des cellules souches hématopoïétiques. Elle constitue donc le lieu de prolifération et de différenciation des lymphocytes B. La moelle osseuse se trouve surtout dans les os plats. Elle est composée de tissus adipeux ou moelle jaune (réservoir de lipides) et de tissus hématopoïétiques ou moelle rouge active renfermant des lignées érythroïdes, myéloïdes et mégakaryocytaires et des nodules lymphoïdes. Les cellules sanguines y sont à différents stades de maturation (tiré de Lord Dubé et l'Italien, 1983).

1.1.6.2 Les ganglions lymphatiques

Ces organes constituent un filtre lymphatique retenant la majorité des antigènes, solubles ou particuliers ou les cellules présentatrices d'antigènes chargées de ceux récupérés dans les tissus périphériques. Ces ganglions comprennent trois zones (tiré de Lord Dubé et l'Italien, 1983):

- La zone corticale (cortex périphérique) contenant les follicules lymphoïdes. L'intérieur de ces follicules nommé centre germinatif, contient des L.B activés par les antigènes mais aussi quelques L.T CD4⁺. Le stroma de ces follicules est composé de cellules folliculaires dendritiques et de macrophages ;
- La zone médullaire (centrale) renfermant des L.B., L.T., macrophages et des plasmocytes;
- La zone para corticale (cortex profond) riche en L.T., contient des L.B.

1.1.6.3 La rate

Parmi ses fonctions, la rate permet la production d'anticorps et de cellules immuno-compétentes contre les antigènes ainsi que l'élimination des particules étrangères grâce à sa population macrophagique (Cordelier D. , 1975). Le parenchyme splénique comprend 3 zones (tiré de Lord Dubé et l'Italien, 1983):

- la pulpe blanche correspondant au tissu lymphoïde. La lymphocytopoïèse se déroule au niveau des follicules lymphoïdes de cette zone;
- la pulpe rouge contient des sinus veineux, des cordons spléniques et des vaisseaux.
- entre les deux zones précédentes, se trouve la zone marginale renfermant des L.T. et des L.B., des hématies, des granulocytes, des plaquettes et des macrophages.

1.1.6.4 Le thymus

Le thymus est le lieu de différenciation et de sélection négative et positive des L.T (tiré de Masmoudi et Ben Ammar, 2000). Les deux lobes du thymus renferment une zone corticale riche en lymphocytes T et une zone médullaire moins riche en lymphocytes.

1.2 Cancer et leucémies

1.2.1 Le cancer

Le cancer est constitué d'une famille diversifiée de maladies complexes caractérisées par des changements génotypiques et phénotypiques (Stratowa t al. , 2001). Différents types de cancers existent. Tous ces types se caractérisent par une croissance désordonnée de cellules anormales et leur propagation dans certaines parties du corps. Un tiers des personnes seront affectées par l'une des formes du cancer (The Norwegian Cancer plan 1997). L'Organisation Mondiale de la Santé (OMS) estime qu'il y aura une augmentation du nombre de décès dus au cancer en Europe et dans le reste du monde et ce à partir de 1996 et jusqu'en

2020. L'augmentation la plus dramatique sera notée dans les pays en voie de développement (WHO 1996).

1.2.2 Les leucémies

Les leucémies sont aussi appelées cancer du sang. Elles sont définies comme des hémopathies malignes (Morse III et al. , 2002) et sont caractérisées par la prolifération anormale et incontrôlée de cellules d'une lignée hématopoïétique qui ne retient pas la capacité de se différencier normalement en cellules sanguines matures (Sawyers et al. , 1991), par leur présence au niveau de la moelle osseuse et dans le sang et généralement par une évolution fatale.

Plusieurs facteurs sont responsables de l'apparition des leucémies dont :

- certaines anomalies chromosomiques congénitales et certains facteurs génétiques. Souvent, des translocations chromosomiques et/ou des inversions sont à l'origine de réarrangements cellulaires ou de la génération de gènes fusionnés codant pour des protéines chimériques (protéines de fusion) (Bernardi et al. , 2002).
- Certains facteurs qui peuvent être le résultat de l'exposition à des radiations, certains produits chimiques, des virus ou des particules.

L'injection de cellules leucémiques à un animal receveur, histocompatible avec l'animal donneur, provoque l'apparition de la leucémie chez le premier, d'où le caractère transplantable des leucémies. Les différents types de leucémies sont définis en fonction du type de cellules sanguines touchées dans le sang et dans la moelle osseuse. Ainsi, une leucémie myéloïde traduit l'apparition des caractères de la lignée myéloïde et une leucémie lymphoïde traduit l'apparition des caractères de la lignée lymphoïde. Lorsqu'une tumeur maligne se présente sous forme solide, elle est qualifiée de lymphome. Expérimentalement, nous avons remarqué que les différents types de leucémies apparaissant chez l'humain, apparaissent aussi chez la souris. Il est donc important de décrire ces différents types pour chacun de ces deux mammifères.

1.2.2.1 Les différents types de leucémies chez l'humain

On distingue deux groupes de leucémies : les leucémies aiguës (myéloblastiques ou AML et lymphoblastiques ou ALL) et les leucémies chroniques (myéloïdes ou CML et lymphoïdes ou CLL).

La leucémie aiguë progressant rapidement, touche des cellules jeunes et empêche la moelle osseuse de fonctionner normalement. Par contre, la leucémie chronique progresse lentement, touche des cellules différenciées et adultes et laisse se former une moelle osseuse normale.

Une nouvelle classification des hémopathies malignes et des leucémies a été récemment proposée par les participants de la communauté internationale (Comité de rédaction de l'OMS), mais les recommandations finales sont apparues dans un fascicule de l'OMS (World Health Organisation Classification of Tumors, International Agency For Research on cancer, Lyon 2001) (www.biocolleges.org/colhemato/Hemato/Whogf.html). C'est la « WHO classification ».

Souhaitée principalement par les cliniciens, dans le but d'une meilleure adaptation aux indications thérapeutiques et tenant compte des syndromes cliniques, cette classification ne bouleverse qu'assez peu les systèmes anciens (Exemples : classification MIC (Morphologique, Immunologique, Cytogénétique) et FAB (essentiellement morphologique)). Elle cherche à compléter les approches classiques de la morphologie microscopique, par les contributions de l'immunologie, de la cytogénétique et de la biologie moléculaire (www.biocolleges.org/colhemato/Hemato/Whogf.html).

La classification préliminaire de l'OMS des désordres hématopoïétiques chez l'humain est présentée ci-dessous (World Health organization, 1997):

- Néoplasmes des cellules B;
- Néoplasmes des cellules T;
- Lymphome de Hodgkin;
- Immunodéficience reliée au désordre lymphoprolifératif;
- Néoplasme histiocytique et cellulaire dendritique et autres désordres associés;

- Leucémies aiguës;
- Syndrome myélodysplasique;
- Désordres chroniques myéloprolifératifs.

1.2.2.2 Les différents types de leucémies chez la souris

Différemment de ce qu'on observe chez l'homme, les leucémies murines sont principalement provoquées par des rétrovirus. Pour la souris, les différents types de néoplasmes hématopoïétiques ont été classés pour la première fois par Dunn en 1954 (Dunn, 1954) sur la base de la cellule d'origine reliée au type de tumeur. La classification la plus récente a été présentée dans la proposition de Bethesda en 2002 (Kogan et al., 2002) et elle est basée sur la lignée cellulaire ayant subi les désordres. Deux sous-groupes ont résulté de cette classification : les désordres lymphoïdes (Morse III et al., 2002) et les désordres non lymphoïdes (Kogan et al., 2002) qui sont eux même divisés en catégories :

Tableau 1.2 Classification des leucémies chez la souris

Les désordres non lymphoïdes	Les désordres lymphoïdes
<ul style="list-style-type: none"> - Leucémie non lymphoïde - Sarcome hématopoïétique non lymphoïde - Dysplasie myéloïde - Prolifération myéloïde (non réactive) 	<ul style="list-style-type: none"> - Néoplasme des cellules B : <ul style="list-style-type: none"> • Néoplasme des précurseurs des cellules B • Néoplasme des cellules B matures - Néoplasme des cellules T : <ul style="list-style-type: none"> • Néoplasme des précurseurs des cellules T • Néoplasme des cellules T matures • Néoplasme des cellules T à caractère indéterminé

De telles classifications uniformes vont permettre de comparer les cancers murins avec les cancers humains (Kogan et al., 2002).

1.3 Les rétrovirus

Les rétrovirus murins sont des virus qui sont capables d'induire la leucémie. Notés MuLV « Murine Leukemia Virus », certains d'entre eux, dont le rétrovirus Moloney (induisant des lymphomes de type T) et le rétrovirus Friend-MLV (impliqué dans la leucémie érythroïde), sont les plus communément étudiés (Coffin et al. , 1997).

Les seuls rétrovirus humains induisant les leucémies sont les virus HTLV-1 « human T-cell leukemia virus type 1 » et HTLV-2 « human T-cell leukemia virus type 2 » (Sawyers et al. , 1991).

Les MuLV se présentent sous deux formes transmises différemment dans la lignée germinale (Sitbon et al. , 2001) : la première forme endogène est verticalement transmise et la seconde exogène est transmise horizontalement et considérée comme infectieuse (découverte par Ludwik Gross lors de l'inoculation de souriceaux d'un jour avec des extraits cellulaires provenant de cellules leucémiques).

1.3.1 Classification des rétrovirus (*Retroviridae*)

En se basant sur les critères de pathogénicité on distingue trois sous-familles (<http://anne.decoستر.free.fr/d1viro/vretrov0.html>):

1.3.1.1 La famille des *Oncovirinae* (*Oncovirus*) est la plus importante des trois familles. Les *Oncovirinae* sont des virus transformants. Ils ont été trouvés chez les insectes, les sangsues, les reptiles, les oiseaux et les mammifères. Les membres de cette famille ont la capacité d'induire le plus grand nombre de tumeurs de différentes origine : développement de tumeurs du tissu conjonctif (sarcomes), de tumeurs des tissus épithéliaux (carcinomes), de leucémies (cas le plus fréquent dans la pathologie rétrovirale) et de lymphomes. Deux rétrovirus appartenant à cette famille et qui ont été découverts par Robert Gallo en 1980 peuvent être cités : le virus HTLV-1 impliqué dans le développement d'une leucémie à lymphocytes T et le virus HTLV-2 qui ne semble être associé qu'à de rares cas de myéloneuropathie primitive proche de la myélopathie associée à HTLV1.

1.3.1.2 Les *Spumavirinae* (*Spumavirus*) doivent leur nom au fait que les cellules infectées *in vitro* présentent des lésions qui ressemblent à de l'écume (*spuma* = mousse). Isolés au hasard au début des années 1950 chez plusieurs espèces animales (singes, bovins, chats, hamsters), ces rétrovirus sont considérés comme des virus "non pathogènes". Ils ont aussi été isolés chez l'homme en 1970.

1.3.1.3 Les *Lentivirinae* (*Lentivirus*) ont été ainsi dénommés car ils sont responsables de maladies à évolution lente (*lentus* = lent). Ce sont des virus cytopathogéniques qui ont d'abord été isolés chez l'animal. Un exemple de rétrovirus appartenant à cette famille est le virus de l'immunodéficience humaine VIH-1 et VIH-2.

1.3.2 Structure des rétrovirus

Les rétrovirus sont des virus enveloppés à développement cellulaire possédant un génome composé d'un dimère d'ARN simple brin de polarité positive, associé à différentes protéines (Sitbon et al. , 2001). Les chaînes d'ARN identiques sont reliées à leur extrémité 5'. Une capsid enveloppe les 2 brins d'ARN et les protéines qui lui sont associées. Ce trio forme la nucléocapside qui est entourée par une enveloppe constituée d'une bicouche lipidique et de 2 types de glycoprotéines : celles transmembranaires et celles de surface. Les protéines de la matrice se situent directement sous l'enveloppe virale. Le schéma du rétrovirus est représenté par les figures 1.2 et 1.3 ci-après.

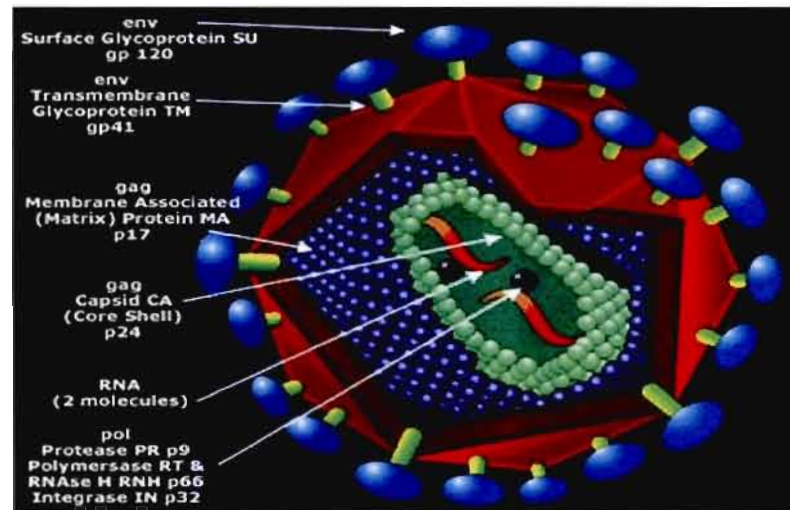


Figure 1.2 Visualisation de la structure du rétrovirus. Les rétrovirus sont constitués des gènes de l'enveloppe (*env*) qui codent pour les glycoprotéines de surface (SU) et pour les glycoprotéines transmembranaires (TM). Ils sont aussi constitués par le gène *gag* codant pour la matrice virale (MA), pour la capside (CA) et pour la nucléocapside (NC). Ces rétrovirus contiennent aussi le gène *pol* codant pour reverse transcriptase inverse (RT) et pour l'intégrase (IN). Deux molécules d'ARN constituent le génome des rétrovirus.

Tiré de http://www.stanford.edu/group/nolan/tutorials/ret_6_gpedesc.html

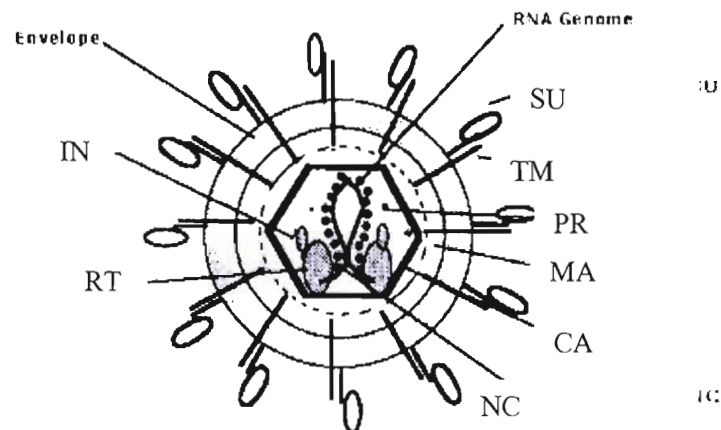


Figure 1.3 Représentation schématique du rétrovirus. SU : glycoprotéines de surface, TM : glycoprotéines transmembranaires, PR : protéase, MA : matrice, CA : capside, NC : nucléocapside, RT : transcriptase inverse, IN : intégrase.

Tiré de http://www.stanford.edu/group/nolan/tutorials/ret_5_struct.html

L'ADN proviral est composé par la succession de trois domaines codants pour les protéines virales GAG, POL et ENV encadrés par les séquences LTR (Long Terminal Repeat) non codantes (Figure 1.4).

Le gène *gag* code pour la synthèse des protéines de la matrice, de la capside et de la nucléocapside. Le gène *pol* code pour les enzymes de la reverse transcriptase inverse et de l'intégrase. Le gène *env*, quant à lui, code pour les glycoprotéines de surface et transmembranaires de l'enveloppe virale. Il existe un autre domaine appelé *pro* codant pour la protéase du virion (Coffin et al. , 1997). Aux extrémités du génome viral, se trouvent trois régions U3, R et U5 constituant le LTR et contenant les signaux de contrôle de l'expression virale (Sitbon et al. , 2001).

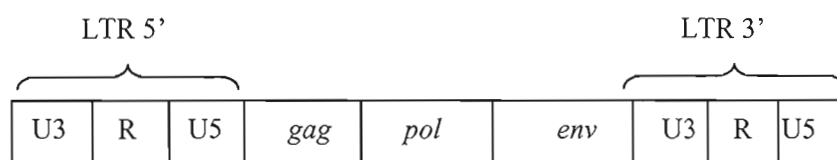


Figure 1.4 Organisation de l'ADN proviral d'un rétrovirus.

Tiré de <http://www.anne.decoستر.free.fr/d1viro/vtelechar/vpoly/hiv05.pdf>

Région R: c'est une petite région de 18 à 250 nt formant une région unique répétée à chacune des extrémités de l'ADN roviral.

U5: cette région est appelée « séquence unique non codante ». Elle mesure entre 75 et 250 nt. Elle représente un promoteur puissant de la transcription.

U3 : cette région est appelée « région unique non codante ». Elle mesure entre 200 et 1200 nt. Elle possède les signaux terminateurs de transcription dont le signal de polyadénylation. Elle constitue aussi un promoteur potentiellement capable d'activer un gène cellulaire situé à proximité.

1.3.3 Cycle de réplication des rétrovirus

Le cycle de réplication des rétrovirus est très particulier (Rassart 2004): le virus est absorbé à l'intérieur de la cellule hôte par attachement de sa glycoprotéine (*env*) de surface aux récepteurs membranaires plasmiques qui lui sont spécifiques. Une fois dans le cytoplasme, la transcriptase inverse produite par le rétrovirus, transcrit l'ARN viral en un ADNc. Ce dernier est constitué de séquences LTR encadrant les gènes codants pour les

protéines virales (Friedman 2002). Acheminé vers le noyau, l'ADNc est inséré dans l'ADN chromosomique à l'aide de l'intégrase virale pour former un provirus stable. A partir de ce moment, les rétrovirus vont détourner en leur faveur, la machinerie transcriptionnelle cellulaire. Etant donné qu'ils sont intégrés, leur génome se comporte comme de l'ADN cellulaire et dépendent donc de l'état de la cellule hôte créant une interaction entre le virus et la cellule. L'insertion des rétrovirus se fait d'une façon aléatoire au niveau de tous les types cellulaires (Sitbon et al. , 2001). Cependant, il semblerait que ces rétrovirus ont des sites d'intégration préférentiels qui sont en faveur de leur prolifération. Parfois, il arrive qu'ils s'intègrent à proximité d'un gène ou d'un proto-oncogène et en altèrent l'expression, ce qui peut être considéré comme un phénomène déclencheur de la cancérisation.

L'ARN polymérase II cellulaire se charge de la transcription du provirus. Non seulement les éléments de régulation contenus dans les LTR du côté 5' ont un rôle très important dans cette transcription mais aussi ceux contenus du côté 3'. En plus de posséder des signaux de polyadénylation définissant la fin du génome rétroviral, le côté 3' peut jouer le rôle d'un promoteur en activant les gènes cellulaires situés en aval. Suite à cette étape de transcription, il y a production d'ARNs représentant à la fois le matériel génétique viral et les ARNm des protéines constituant le rétrovirus. Ces deux produits seront assemblés en particules virales au niveau de la membrane plasmique de la cellule et sont libérés par bourgeonnement pour former des virions matures qui à leur tour, vont infecter d'autres cellules (Figure 1.5).

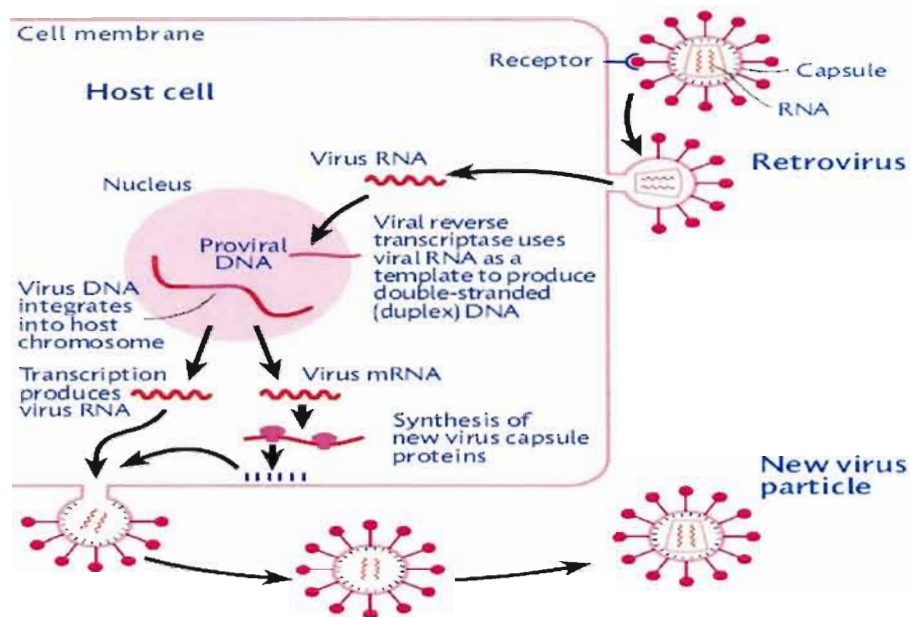


Figure 1.5 Mécanisme de répllication des rétrovirus.

Tiré de <http://www.bioscience-explained.org/EN1.2/pdf/virusEN.pdf>

1.3.4 Les oncogènes

Les oncogènes ont été classiquement définis comme des gènes à caractère dominant. Le dérèglement de leur expression ou l'altération de leur structure, contribue au phénotype transformé d'une cellule (Gisselbrecht 2003). En 2003, 100 oncogènes codant pour des oncoprotéines possédant différentes fonctions, ont été associés aux hémopathies malignes humaines.

En fonction des oncoprotéines pour lesquelles ils codent, les oncogènes sont répartis en 8 classes majoritaires (25.<http://www.chups.jussieu.fr/polys/cancero/POLY.Chp.3.5.html>) :

- les facteurs de croissance assurant une boucle de régulation autocrine. Comme exemple, nous pouvons citer les proto-oncogènes codants pour les protéines de la famille des «Fibroblast Growth Factors» (FGF) ;

- les récepteurs transmembranaires de facteurs de croissance tels que le proto-oncogène *erb B* codant pour le récepteur à l'«Epidermal Growth Factor» (EGF) ;
- les G-protéines ou protéines membranaires liant le GTP tels que les proto-oncogènes de la famille *ras* ;
- les protéines tyrosines-kinases membranaires ;
- les kinases cytosoliques ;
- les récepteurs intra-membranaires tel que le proto-oncogène *erbA* codant pour le récepteur aux hormones thyroïdiennes
- les protéines à activité nucléaire qui en interagissant avec l'ADN, sont capables de contrôler la transcription de gènes cibles. Comme exemple, nous pouvons citer, les proto-oncogènes *fos*, *jun* et *c-myc*.

Les anti-oncogènes sont les gènes suppresseurs de tumeurs. Lorsqu'ils sont de nature protéique, ces anti-oncogènes sont synthétisés au cours de l'expression d'un gène qui joue donc un rôle antagoniste de l'oncogène. La perte de fonction d'un anti-oncogène est responsable de la non-production de l'inhibiteur d'un oncogène aboutissant au processus tumoral. Parmi ces anti-oncogènes, nous pouvons citer p53, P105 (ou pRb), p107, p13,...

1.3.5 Activation des oncogènes

On distingue deux types de rétrovirus: les virus non défectifs et les virus défectifs (Rassart 2004) :

- les virus défectifs contiennent un oncogène d'origine cellulaire remplaçant leur gène de structure les empêchant de se répliquer d'une manière autonome au niveau de la cellule hôte. Pour se répliquer, ce manque doit obligatoirement être comblé par la présence d'un rétrovirus auxiliaire. Ces virus sont à transformation rapide (entre 2 à 3 semaines post-infection).
- les virus non défectifs n'ont pas besoin de la présence d'un autre rétrovirus pour se répliquer au niveau de la cellule hôte. Leur génome ne porte pas d'oncogène d'origine cellulaire à la place de leurs gènes de structure. Ainsi pour provoquer la

formation de tumeurs, ils doivent s'intégrer à proximité d'un proto-oncogène cellulaire. Ces virus sont à transformation lente (3 à 9 mois).

Pour les MuLV, la mutagénèse insertionnelle constitue le mécanisme principal de l'oncogénèse par activation ou inactivation des gènes cellulaires (Denicourt et al. , 1999). Les proto-oncogènes peuvent être activés par les rétrovirus non défectifs de différentes manières (Coffin et al. , 1997) :

- par insertion de la séquence "enhancer" du LTR activatrice du promoteur cellulaire. Le rétrovirus n'a qu'à s'intégrer à distance avant ou après un proto-oncogène et dans les deux orientations possibles par rapport à la transcription.
- par insertion du rétrovirus à l'intérieur d'un proto-oncogène au niveau d'un exon entraînant la mutation de celui-ci ou à l'extrémité 3' du gène réarrangé.
- par insertion d'un promoteur LTR viral nécessitant l'intégration du virus au niveau du proto-oncogène à son extrémité 5' ou dans un intron.

1.3.6 Oncogènes et leucémies humaines

C'est au début des années 1970 qu'a eu lieu, en cytogénétique, l'introduction des techniques de bandes chromosomiques (Gisselbrecht 2003). Ces techniques ont permis d'identifier chacun des chromosomes d'une façon très précise ainsi que de caractériser les remaniements chromosomiques qui se trouvent au niveau des cellules leucémiques. Soixante-cinq pourcent des cellules leucémiques humaines présentent un remaniement chromosomique qui est le plus souvent unique (Look 1997). Les anomalies chromosomiques les plus fréquentes

sont reliées à des inversions intra-chromosomiques ou à des translocations lorsque deux chromosomes s'échangent du matériel génétique. Deux découvertes, datant de 1982, conduisirent à rechercher systématiquement des altérations chromosomiques dans les hémopathies malignes humaines. Parmi ces découvertes, nous pouvons citer l'identification du gène MYC (localisé chez l'homme à la région

chromosomique 8q24 et se trouvant dans plusieurs modèles de leucémies lymphoïdes expérimentales) se présentant comme une cible des translocations du lymphome de Burkitt (Dalla-Favera et al. , 1982 ; Taub et al. , 1982).

1.3.7 Le rétrovirus murin Graffi

C'est en 1957 qu'Arnold Graffi a pu isoler le rétrovirus murin Graffi à partir d'un filtrat acellulaire provenant du sarcome d'Ehrlich. Il a montré que ce rétrovirus murin non déficient induisait principalement des leucémies myéloïdes. Graffi et son équipe ont été confrontés au fait que les leucémies induites par ce rétrovirus n'étaient pas uniquement de type myéloïde et que certains autres types de leucémies pouvaient apparaître après plusieurs passages (Graffi et Schramm 1966 ; Fey et Graffi 1965; Ru et al. , 1993). Ils ont ainsi considéré ce phénomène comme étant une diversification hématologique. Ensuite, en 1993 et grâce à des outils de clonage moléculaire, dans le laboratoire du docteur Eric Rassart, Ru et al. (1993) ont démontré que le filtrat acellulaire contenant le rétrovirus murin Graffi était en fait constitué de deux variants viraux (GV-1.2 et GV-1.4). L'équipe du docteur Eric Rassart a aussi noté que les leucémies myéloïdes induites par le rétrovirus murin Graffi présentaient parfois des réarrangements au niveau des chaînes lourdes d'immunoglobulines (JH), caractéristiques des lymphocytes B ou bien au niveau des récepteurs TCR, caractéristiques des lymphocytes T et ont défini cela comme étant une « infidélité de linéage ». Dernièrement, Voisin et al. (2006) ont démontré qu'en fait le rétrovirus murin Graffi est un virus très polyvalent qui induit différents types de leucémies (Voisin et al. , 2006) dont les lymphoïdes (de type B et de type T) et non lymphoïdes (érythroïdes, mégacaryocytes et myéloïdes). Cette grande variété de leucémies induites semble liée à la souche de souris (Balb/c, NFS ou FVB) et au variant rétroviral (GV1.2 ou GV1.4).

Trente tumeurs induites par le rétrovirus murin Graffi ont été étudiées par l'équipe du Dr. Eric Rassart dans le but de comprendre les mécanismes de la leucémogénèse ainsi que les gènes touchés et le rôle du MuLV Graffi dans cette transformation (Denicourt et al. , 1999). Parmi ces 30 tumeurs, 6 étaient réarrangées pour le gène *c-myc*, 1 pour le gène *Spi-1/Pu-1*, 3 pour le gène *Fli-1* et 1 tumeur était réarrangée pour les gènes *Pim-1* et *c-myc*. Ces

données permettent de conclure que *c-myc* est le gène fréquemment réarrangé et se comparent aux leucémies lymphoïdes de type B et T induites par le Mo-MLV ou le Fr-MLV (Selten et al. , 1984 ; Corcoran et al. , 1984) et aussi dans le cas de leucémies myéloïdes aiguës (Strassburg et al. , 1992).

Les pourcentages relativement faibles du réarrangement des gènes *c-Myc* (20%), *Fli-1* (10%), *Pim-1* (3.3%) et *Spi-1/Pu-1* (3.3%) montrent qu'il existe encore d'autres proto-oncogènes cellulaires pouvant être impliqués dans la transformation par le rétrovirus Graffi. En effet, dernièrement un nouveau site commun d'intégration nommé *Gris-1* (pour Graffi intégration site 1) a été découvert dans les leucémies induites par Graffi. L'intégration d'un rétrovirus à son niveau provoque l'augmentation de l'expression d'une Cycline D2 tronquée (Denicourt et al. , 2003).

Toujours dans le but de mieux caractériser le rétrovirus murin Graffi, d'autres recherches présentées en annexe ont été effectuées (Voisin et al. , 2006 ; Annexe A). Ces résultats montrent bien que la souche de souris NFS infectée par le variant moléculaire 1.4 du rétrovirus Graffi est celle qui donne le plus de variations au niveau du type de leucémies.

1.4 Les micropuces à ADN : Méthode d'analyse de l'expression génique des cellules de souris leucémiques et non leucémiques

Cette révolution génomique offre de nouveaux outils pour l'étude de processus biologiques et complexes à l'échelle pan-génomique. La méthode des micropuces à ADN permet l'étude simultanée de l'expression de milliers de gènes et l'obtention d'une vue globale de l'expression génique (Deyholos et Galbraith, 2001 ; Geiss et al. , 2001). Un grand nombre de nouvelles applications est offert par cette technique. Ces applications sont, à la fois, fondamentales et cliniques. Les résultats obtenus par cette technique amènent les chercheurs à se poser de nouvelles questions afin d'essayer d'élucider certains mécanismes moléculaires complexes. Depuis longtemps, les biologistes utilisent des techniques de base afin de disséquer les phénomènes biologiques (Northern blots, Southern blots, ...). Ces

techniques traditionnelles ne peuvent que résoudre des points très spécifiques témoignant de leur limitation.

Afin d'étudier l'expression génique des cellules leucémiques et dans le but de déterminer les gènes fortement impliqués dans l'apparition et dans l'évolution du cancer, la technique de microarrays ou micropuce à ADN du système Affymetrix, s'avère d'un grand intérêt.

Le principe de cette méthode est le suivant : la puce à ADN, proposée par la société Affymetrix, est une lame de verre qui contient certaines séquences représentatives des gènes de souris qui ont été déjà identifiés et où chaque gène est représenté par des oligonucléotides synthétisés directement sur un substrat solide par photolithographie. Chaque gène présent sur une puce à ADN destinée à des études d'expression, est représenté par un ensemble d'oligonucléotides mimant la séquence du gène. Ces oligonucléotides sont souvent choisis dans sa région 3', réduisant ainsi les risques d'hybridations croisées avec des séquences homologues de ce gène. Ces oligonucléotides sont divisés en une série de 20 oligonucléotides qui sont homologues aux ARNm de la souris (notés séquences PM pour Perfect Match) et de 20 oligonucléotides (notés séquences MM pour MisMatch) qui sont identiques aux séquences PM mais comportant le changement d'une seule base par son complémentaire au niveau de la position centrale 13. Ceci permet la détection des hybridations non spécifiques et aussi la quantification des ARNm fortement exprimés.

La puce à ADN est d'abord hybridée avec des ARNm de tumeurs (provenant des cellules déjà triées) après qu'ils aient été convertis en ADN complémentaires (ADNc), qui à leur tour sont convertis en ARNc biotinylés. Ces derniers sont marqués avec un fluorochrome couplé à la streptavidine. Ensuite, la puce ou « chip » est scannée et l'hybridation des ARNc avec les oligonucléotides est détectée par l'émission d'une fluorescence dont l'intensité émise est proportionnelle au niveau d'expression de chaque gène (elle est d'autant plus forte qu'il y a de copies d'ARNm correspondant à chaque gène) (figure 1.6).

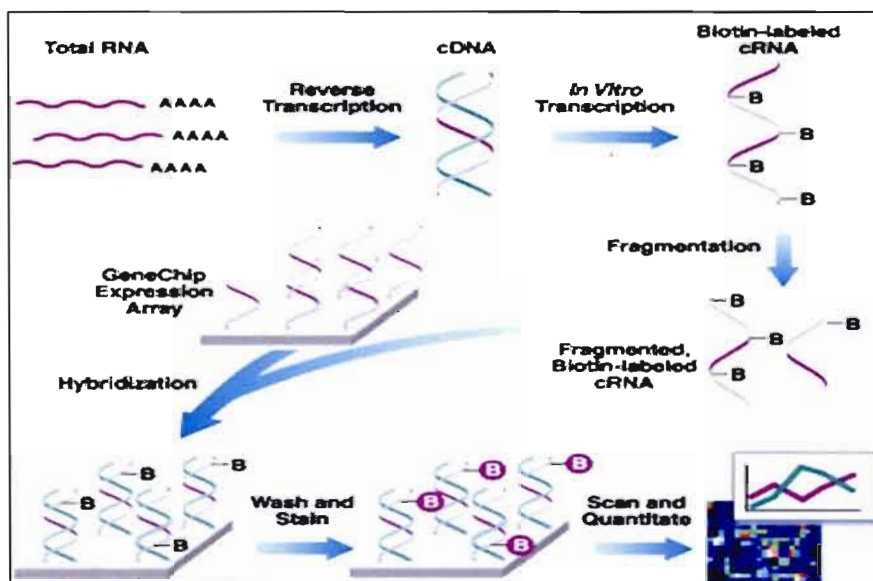


Figure 1.6 Principe de la méthode des micropuces à ADN « Microarrays ».

Tiré de http://www.affymetrix.com/technology/ge_analysis/index.affx

1.5 Hypothèse de travail

Le rétrovirus murin Graffi est un virus très polyvalent qui a la capacité d'induire différents types de leucémies. Il est constitué de deux variants rétroviraux (GV-1.2 et GV-1.4). Ainsi, en injectant l'un ou l'autre des clones de ce virus à différentes souches de souris nouveau-nées (Balb/c, NFS et FVB), nous pouvons nous attendre à obtenir des leucémies lymphoïdes de type T et B et des leucémies non lymphoïdes de type myéloïdes, érythroïdes et mégakaryocytaires. L'obtention d'un certain pourcentage de chaque type devrait dépendre de la souche de souris et du variant rétroviral utilisés. La souche NFS injectée avec le clone moléculaire GV-1.4 devrait permettre l'obtention de plusieurs types de leucémies comme cela a été montré dans l'annexe A (résultats obtenus par Véronique Voisin). Il est connu que chaque lignée hématopoïétique exprime des marqueurs de surface qui lui sont bien spécifiques. Il serait donc possible de déterminer le type de leucémies induites par le rétrovirus chez chaque souris par la méthode de la cytométrie de flux et à l'aide d'anticorps reconnaissant certains marqueurs de surface. Il serait aussi possible, à partir de différentes

tumeurs, d'isoler une population leucémique pure et homogène par la technique du tri cellulaire par les billes magnétiques.

Le but ultime de ce projet est l'analyse de l'expression génique des cellules leucémiques lymphoïdes T et B par la technique des micropuces à ADN pour mieux comprendre les mécanismes qui provoquent la dérégulation d'une cellule hématopoïétique. En utilisant cette technique, sur les ARN triés par la technique des billes magnétiques, nous devrions être capable d'étudier l'expression de plusieurs milliers de gènes à la fois. Nous nous attendons à voir apparaître des gènes fortement exprimés dans un seul type de leucémie (T ou B). Il se pourrait donc que ces gènes soient spécifiques au type de leucémie correspondant. Toutefois, cette différence pourrait aussi être reliée à une différence de linéage. Nous nous attendons aussi à récupérer des gènes qui se trouvent fortement exprimés dans les deux types de leucémies. Grâce à ces gènes, considérés comme communs aux leucémies T et B, il serait possible de classer les leucémies.

Finalement cette technique pourrait mener à l'identification des changements au niveau de l'expression de certains gènes, ce qui pourrait être utilisé pour découvrir de nouveaux bio-marqueurs ainsi que pour essayer de mieux comprendre le mécanisme d'induction de la leucémie.

CHAPITRE II

MATÉRIEL ET MÉTHODES

2.1 Rétrovirus murin Graffi

Des cellules NIH3T3 en culture, infectées chroniquement par l'un des deux clones moléculaires GV-1.4 et GV-1.2 du rétrovirus murin Graffi, ont été mises en culture. Le surnageant de ces cellules, récupéré selon la méthode de « Rapid Harvest », contient les particules virales correspondant aux deux variants rétroviraux.

Les trois souches de souris que nous avons choisies sont les Balb/c, les NFS et les FVB/n auxquelles 100 μ l (le titre viral est de 3.10^6 PFU pour GV-1.2 et de 1.10^6 PFU pour GV-1.4) de ce surnageant sont injectés par voie intra péritonéale à la naissance. L'âge de ces souriceaux ne doit pas excéder les 2 jours. La leucémie survient alors après une période de latence variable de quelques mois.

2.2 Les organes tumoraux

Les souris ayant atteint le dernier stade de la maladie sont euthanasiées et leurs organes hématopoïétiques tels que la rate, le thymus (lors d'un thymome), la moelle osseuse (à partir des os du fémur), le sang, ainsi que les ganglions (lorsqu'il y en a), sont extraits. D'autres organes sont aussi extraits comme le foie, les reins et le cerveau pour servir de tissus témoins. Une partie des organes (rate, thymus, ganglions, foie, reins et cerveau) est congelée à -80°C pour une utilisation ultérieure. L'autre partie (rate, foie, moelle osseuse, thymus, ganglions, cellules sanguines) peut être utilisée directement. Les organes à analyser par cytométrie de flux et par tri sont choisis selon l'importance de la population leucémique qui s'y trouve. Cette importance est déterminée à l'autopsie de la souris. La même procédure a été appliquée pour les souris contrôles de l'ensemble des 3 souches Balb/c, NFS/n et FVB (Voir les tableaux 2.2 et 3.1 contenant les différentes souris contrôles analysées). Cependant,

les souris adultes montrent une involution de la taille du thymus. Pour cette raison, nous choisissons de sacrifier de jeunes souris contrôles dont l'âge ne dépasse pas les trois mois.

2.3 Les cellules leucémiques

♦ Préparation des cellules pour analyse par cytométrie de flux

La rate, le foie, le thymus, la moelle osseuse et les ganglions sont placés dans un milieu contenant du PBS (Phosphate Buffered Saline), 2% de FCS (Sérum de veau foetal) (pour éviter une internalisation des récepteurs de surface) et 0.1% d'azide de sodium et les cellules sont séparées par broyage mécanique. Afin d'éliminer les érythrocytes, une exposition des cellules sanguines et des cellules de la rate à une solution de lyse (constituée de 80g de NH_4Cl (Chlorure d'ammonium), 10g de KHCO_3 (Bicarbonate de potassium), de 3.7g de Na_4EDTA (Ethylenediamine Tetraacetic Acid tétrasodique) et de 1000ml d'eau distillée) suivie d'une incubation sur glace pendant 5 minutes, sont nécessaires. Pour la rate et le sang, une première étape de centrifugation effectuée pendant 5 min à 500 x g suivie d'un lavage avec du PBS et d'une deuxième centrifugation pendant 3 min à 500 x g permettent de récupérer un culot de cellules vivantes. Le foie, le thymus, la moelle osseuse et les ganglions sont soumis à une seule étape de centrifugation pendant 5 min à 500 x g.

L'ensemble des cellules provenant de tous ces organes est récupéré dans une solution contenant du PBS, 2 % de FCS et 0.1 % d'azide de sodium.

♦ Préparation des cellules pour trier la population cellulaire désirée

Le même protocole utilisé pour le foie, le thymus, la moelle osseuse et les ganglions est suivi.

2.4 La cytométrie de flux (FACScan)

L'intérêt de cette méthode utilisant l'appareil du FACScan (Becton Dickinson), est de distinguer et de séparer différentes populations cellulaires d'une même suspension

cellulaire. Plusieurs paramètres peuvent être étudiés dont la taille, la granulation et la fluorescence de chaque cellule.

Protocole

Les cellules présentes dans chacune des suspensions cellulaires provenant des différents organes tumoraux, sont réparties dans une série de tubes (10^6 cellules dans chacun). Dans ces tubes, les cellules sont marquées avec 0.4 μ l d'anticorps (pour 10^6 cellules), dilués au $1/125^{\text{ème}}$, conjugués avec l'un ou les deux fluorochromes suivants : la fluorescéine (FITC) et la phycoérythrine (PE).

Le tableau 2.1 ci-après précise la liste des anticorps utilisés de la compagnie BD Biosciences Pharmingen :

Tableau 2.1 Différents anticorps utilisés pour l'immunophénotypage des leucémies par cytométrie de flux et pour le tri des cellules par les billes magnétiques

Anticorps	Spécificité
Mac-1	Macrophages
Gr-1	Granulocytes
TER-119	Erythrocytes
CD71	Erythrocytes
CD3	Lymphocytes T
CD90	Lymphocytes T
CD4	Lymphocytes T
CD8	Lymphocytes T
CD19	Lymphocytes B
CD45	Lymphocytes B
Sca-1	Cellules souches
CD34	Progéniteurs
C-Kit	Cellules souches
CD41	Mégakaryocytes
CD11c	Macrophages
DX5	Cellules NK (Natural Killer)
IP (Iodure de Propidium)	Cellules mortes

Pour éliminer les éventuelles fixations non spécifiques des différents anticorps employés, il est important de mentionner que des isotypes (anticorps « non réactifs ») sont aussi utilisés.

Pour fixer chaque anticorps à sa molécule de surface spécifique, une incubation de 40 min sur glace et dans l'obscurité est nécessaire, suivie d'un lavage au PBS et d'une centrifugation à 10 °C pendant 5min et à 500 x g. Enfin les cellules marquées avec les anticorps correspondant sont resuspendues dans du PBS pour être analysées par FACScan.

2.5 Le tri des cellules par la technique des billes magnétiques

Les cellules présentes dans chacune des suspensions cellulaires provenant des organes tumoraux choisis sont incubées sur glace et à l'obscurité pendant 40 min, avec l'anticorps spécifique de la lignée leucémique. Nous avons adapté la quantité d'anticorps au nombre de cellules à trier pour obtenir un minimum de 10^7 cellules après tri. Après un lavage au PBS et une centrifugation pendant 6 min à 500 x g, toutes les cellules, dont celles qui ont fixé les anticorps, sont resuspendues dans une solution de PBS et dans une solution de PBS, 2 % de FCS et 0.1 % d'azide de sodium.

La première étape du tri consiste à rajouter 80 µl de cocktail d'anticorps, dilués 12.5 fois, aux cellules préalablement marquées par l'anticorps spécifique de la population leucémique à récupérer. Par exemple, si les résultats obtenus par cytométrie de flux montrent que notre leucémie est lymphoïde de sous-type $CD4^+/CD8^-$, nous trions notre population avec l'anticorps CD4. L'ensemble des cellules marquées et du cocktail d'anticorps est incubé une première fois pendant 15 min à température de la pièce puis une deuxième fois. Par la suite 40 µl de billes magnétiques, dilués au $1/27^{ème}$, sont rajoutés. Ces billes vont reconnaître et fixer les anticorps contenus dans le cocktail d'anticorps. Pour récupérer la population leucémique désirée, et pour augmenter le volume de notre échantillon, le tube contenant tout le mélange est ajusté à un volume de 2.5 ml avec du PBS. Ce tube est par la suite mis en présence d'aimant pendant 5 minutes. C'est ainsi que les billes magnétiques qui sont fixées à la population que nous voulons récupérer vont se coller aux parois du tube. Quatre ou cinq lavages avec du PBS sont nécessaires pour se débarrasser des populations cellulaires non désirées qui sont présentes dans le surnageant. Un pourcentage de pureté supérieur ou égal à 90% est obtenu suite à ces lavages. Une fois le tri effectué, les cellules contenues dans le PBS sont centrifugées à 2500 x g pendant 5 minutes à 4 °C. Finalement, pour extraire l'ARN à partir de ces cellules, le culot est repris dans 1ml de Trizol (méthode du Trizol® (Invitrogen- Life Technologies) puis congelé à -80 °C.

2.6 Analyses en biologie moléculaire

2.6.1 Etude de l'expression génique de quelques gènes par Northern Blot

L'ARN est extrait par le Trizol selon la méthode suivante : un morceau de rate (environ 100 ng) est mis en présence d'1 ml de Trizol. L'organe est ensuite broyé grâce à l'appareil Utraturax pour obtenir un mélange homogène. On y rajoute un même volume de chloroforme suivi d'une incubation de 2 min entre 15 et 30 °C et d'une centrifugation pendant 15 min à 12000 x g et à 4 °C. La phase aqueuse contenant l'ARN cellulaire, est incubée dans l'isopropanol pendant 10 min entre 15 et 30 °C. L'ARN est ensuite précipité après centrifugation pendant 10 min à 12000 x g et à 4 °C. Le culot d'ARN est lavé avec de l'éthanol 75%. L'ARN ainsi purifié est centrifugé pendant 5 min à 7500 x g et à 4 °C. Il est enfin redissout dans de l'eau («RNase-free water») et incubé pendant 10 min à 55 °C. La qualité de cet ARN est vérifiée au moyen d'un gel d'agarose à 0.8% dénaturant en présence de formaldéhyde. Il est ensuite transféré en conditions dénaturantes sur membrane de nylon (Fisher Scientific Co) et hybridé avec les sondes radioactives spécifiques d'une lignée cellulaire hématopoïétique donnée de la même manière que cité ci-dessous afin d'analyser l'expression des gènes correspondants.

2.6.2 Hybridation de la sonde radioactive pour le Northern Blot

La membrane contenant l'ARN transféré est pré-hybridée au minimum pendant 30 minutes à 42 °C dans du tampon constitué de formamide 50%, de 20X SSPE (« Saline Sodium Phosphate EDTA ») (3 M chlorure de sodium (NaCl), 0.2 M phosphate de sodium monobasique (NaH_2PO_4)), de Denhardt's 100 X (2% Ficoll, 2% tolyvinylpyrrolidone, 2% bovine sérum albumine (BSA)) et de SDS (Sodium dodecyl sulfate) 10%. Suite à cette première incubation, la membrane est hybridée pendant un minimum de 15 minutes à 42 °C avec un tampon d'hybridation constitué de formamide 50%, de 20 X SSPE, de Denhardt's 100x, de SDS 10% et Dextran sulfate 50%. Après rajout de la sonde dénaturée, l'hybridation se fait pendant toute la nuit à 42 °C. Deux lavages de la membrane de 15 minutes sont effectués à 42 °C avec une solution de 2SSC (« salin citrate de sodium ») (3M NaCl, 0.3 M

sodium citrate). Un dernier lavage effectué pendant un minimum de 15 minutes à 65°C dans une solution de 2% SSC- 0.1% SDS est nécessaire pour éliminer les sondes non fixées à la membrane étant donné que c'est la solution de lavage la plus astringente. Finalement, la membrane est exposée à un écran PhosphorImager (Kodak).

2.7 Les micropuces : Méthode d'analyse de l'expression génique des cellules leucémiques et non leucémiques

En raison de l'inexistence de l'équipement nécessaire, cette partie n'a pas pu être réalisée à l'UQAM. Elle a été réalisée au Centre Génomique de Montréal (à l'Université de McGill) en collaboration avec le docteur Tom Hudson.

La compagnie Affymetrix fournit un « set » nommé « The Genechip® Mouse Genome 430 2.0 Array set » qui permet d'analyser plus de 39000 transcrits par puce. Toutes les informations concernant les séquences utilisées par Affymetrix dérivent du GenBank®, du dbEST et du Ref Seq. Les séquences clusters ont été créées à partir de la banque de données UniGene (Build 107, June 2002) et des séquences du génome de la souris à partir du Whitehead Institute for Genome Research (MGSC, April 2002). Parmi les 39000 transcrits se trouvent plus de 34000 gènes murins qui ont été bien caractérisés sur le plan biochimique et moléculaire. Ainsi, la compagnie Affymetrix a choisi des gènes de séquences connues ainsi que des EST (« Expressed sequence tag ») et a créé des « ProbeSets » pour chaque transcrit (une probeset est représentée par une série de 11 à 20 oligonucléotides représentant un gène donné). Pour cette raison, le nombre total de « ProbeSets » représentés sur une micropuce à ADN est supérieur à 45000 (cela veut dire, que par gène, plusieurs « ProbeSets » sont utilisées pour le même transcrit dans le but, d'une part de servir comme un moyen pour confirmer le niveau d'expression d'un gène représenté par plus d'une « probeSet » et d'autre part pour identifier de nouveaux transcrits épissés ou bien des transcrits épissés déjà connus). Deux logiciels sont utilisés : le « GeneChip Scanner 3000 » qui permet une haute résolution de scannage et le « GeneChip Operating Software (GCOS) v1.1.1 » (qui inclut le « GeneChip Scanner 3000 High-Resolution Scanning Patch »).

Le protocole décrit par Affymetrix, nécessite une quantité d'ARN supérieure ou égale à 1 µg. Suite à l'étape de purification de l'ARN extrait à partir des cellules que nous avons trié, ce dernier est envoyé au Centre Génomique de Montréal où il a été converti en ADN complémentaire et biotinylé pour servir comme sonde sur les micropuces (Voir la figure 2.1 ci-dessous).

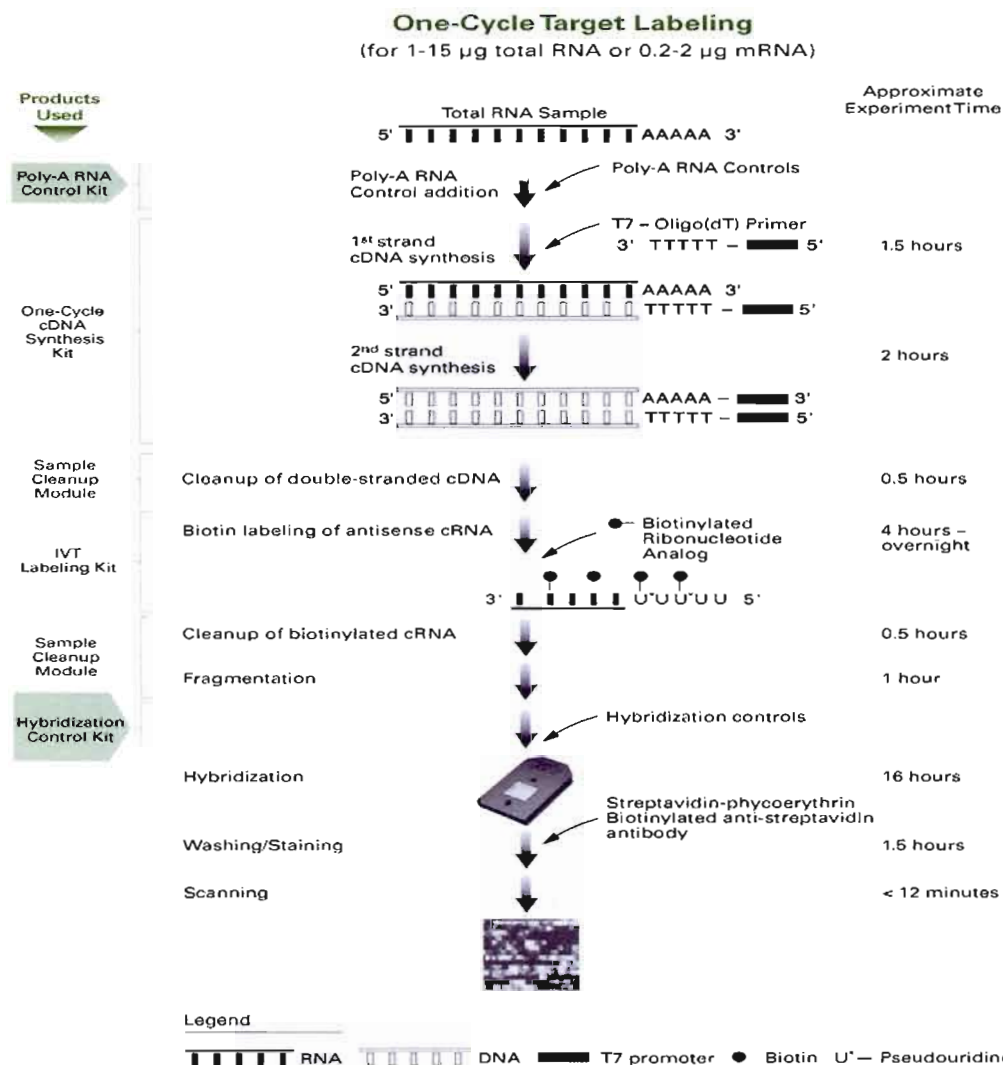


Figure 2.1 Schématisation des différentes étapes nécessaires pour la préparation et l'hybridation des fragments d'ARN sur les micropuces à ADN. Ce protocole est réalisé pour une quantité d'ARN qui varie entre 1 et 15 µg d'ARN total ou bien entre 0.2 et 2 µg d'ARNm. Tiré de : http://www.affymetrix.com/products/reagents/specific/one_cycle_target_control.affx

Pour pouvoir comparer les deux types de leucémie (T et B) entre elles, nous avons utilisé trois « puces », une « puce » par souris et par leucémie selon le tableau 2.2 de planification suivant :

Tableau 2.2 Caractéristiques phénotypiques des échantillons analysés par les micropuces à ADN

Type de cellules	Type de leucémies	Nombre de « chip »
Lymphocytes T leucémiques	CD4+ CD8-	3 « puces » (1 par souris)
	CD8+ CD4-	
	CD4+ CD8+	
Lymphocytes B leucémiques	CD45R low / CD19+	3 « puces » (1 par souris)
	CD45R+ / CD19+	
	CD45R+ / CD19+ / SCA1+	
Lymphocytes T non leucémiques	CD4+ CD8+	1 « puce »
Lymphocytes B non leucémiques	CD45R+ / CD19+	1 « puce »

Nous avons envoyé l'ensemble de ces échantillons au Centre Génomique de Montréal dans le but d'étudier l'expression différentielle des gènes au niveau des leucémies induites par le rétrovirus murin Graffi.

Pour des raisons budgétaires, il nous était impossible d'utiliser trois puces pour chacune des trois leucémies T et pour chacune des trois leucémies B d'autant plus que ce projet est co-réalisé avec Véronique Voisin qui elle-même a utilisé trois puces pour chacune des leucémies érythroïdes, mégakaryocytaires et une puce pour une leucémie myéloïde en plus d'une puce contrôle pour les trois leucémies érythroïdes.

Il est important de noter que de très nombreuses données sont obtenues après analyse par la méthode des microarrays. Il a été donc nécessaire de trier toutes ces informations. A cet effet, plusieurs méthodes sont disponibles dont le classement hiérarchique en clusters qui a pour but, entre autres, de regrouper les gènes exprimés de la même manière dans différents types de leucémies. Nous avons eu recours, entre autres, à un programme d'analyse statistique nommé RMA (Robust multichip average) afin de mieux nous assurer de la validité de nos résultats.

2.8 RMA et sélection des gènes

L'analyse des résultats obtenus a été réalisée à l'aide d'un outil informatique nommé RMA. Ce dernier nous a permis d'évaluer le ratio du log entre les cellules leucémiques T et B et leur contrôle respectif. En d'autres termes, pour chaque gène, Affymetrix nous présente le niveau de son expression sous la forme d'un nombre, logarithmique (voir l'exemple dans tableau ci-dessous). Grâce à ce nombre il est possible de savoir si un gène est faiblement ou fortement exprimé dans une leucémie par rapport à son contrôle en appliquant la formule suivante :

$$\log A - \log B = \log (A/B)$$

Exemple

Nos résultats ont montré que les valeurs logarithmiques d'expression du gène *Mettl1* (methyltransferase-like 1) étaient de 4.12 pour le contrôle T et de 6.46 pour la leucémie T1. Afin de déterminer si le gène *Mettl1* est sur ou sous exprimé dans les leucémies T (T1) par rapport au contrôle T (CT) et aussi pour déterminer la valeur de cette sous ou sur expression, nous appliquons la formule précédente. Cela donne :

$$T1 - CT = \log (T1 / CT)$$

$$\text{Or } T1 - CT = 6.461847154 - 4.128692913 = 2.333154241$$

$$\text{Cela revient à dire que : } \log (T1 / CT) = 2.333154241$$

$$\text{Donc } T1 / CT = 5.04$$

Ainsi, le gène *Mettl1* est environ 5 fois plus exprimée (5.04 fois) dans la leucémie T1 par rapport au contrôle T (CT).

Au cours de la première sélection et en appliquant le raisonnement précédent :

- nous avons choisi les gènes variant au minimum de 4 fois entre les 3 leucémies T par rapport au contrôle T. Ceux-ci ne doivent pas varier plus de 1.5 fois entre les 3 leucémies B par rapport au contrôle B. Ces gènes ont été considérés comme étant spécifiques aux leucémies T;
- nous avons choisi les gènes variant au minimum de 4 fois entre les 3 leucémies B par rapport au contrôle B. Ceux-ci ne doivent pas varier plus de 1.5 fois entre les 3

leucémies T par rapport au contrôle T. Ces gènes ont été ainsi considérés comme étant spécifiques aux leucémies B.

- nous avons choisi les gènes variant au minimum de 4 fois entre les 3 leucémies T par rapport au contrôle T et aussi entre les 3 leucémies B par rapport au contrôle B. Ces gènes seraient communs aux leucémies T et B.

Suite à cette première sélection, nous avons réalisé une recherche bibliographique dans le but de trouver des gènes :

- qui seraient impliqués dans d'autres types de cancers mais qui ne le seraient pas encore dans les leucémies;
- qui n'auraient jamais été associés ni aux cancers ni aux leucémies mais qui seraient nécessaires au bon fonctionnement de l'organisme;
- auxquels aucune fonction n'ait été associée.

2.9 Sélection de quelques gènes pour confirmer le résultat obtenu par les micropuces à ADN

Afin de valider le résultat obtenu suite à l'analyse par la technique des micropuces à ADN, nous avons sélectionné 7 gènes (Tableau 2.3). Ces derniers ont été choisis à partir de l'ensemble des gènes issus de la recherche bibliographique. C'est ainsi que nous nous sommes intéressés à ceux dont l'expression se trouve plus spécifiquement dérégulée au niveau des leucémies lymphoïdes. En d'autres termes, les leucémies lymphoïdes et les leucémies non lymphoïdes peuvent exprimer les mêmes gènes puisqu'ils proviennent de la même cellule souche. Nous avons donc amélioré nos paramètres de sélection et notre stratégie de telle sorte que la dérégulation de l'expression des gènes sélectionnés soit le plus que possible spécifique aux leucémies lymphoïdes et ne soit pas commune avec les leucémies non lymphoïdes. Ceci a été possible à réaliser étant donné que cette étude a été réalisée en parallèle avec l'étude des leucémies non lymphoïdes (menée par Véronique Voisin). Ainsi, pour sélectionner ces gènes, nous avons calculé la moyenne d'expression de chaque gène au niveau de tous les échantillons (leucémies et contrôles). Ces échantillons sont constitués des 3 leucémies T, des 3 leucémies B, des contrôles T et B et aussi de ceux qui font partie du projet

de Véronique Voisin soit, une leucémie myéloïde, 3 leucémies mégakaryocytaires, 3 leucémies érythroïdes et un contrôle érythroïde. Par la suite, pour chacun de ces gènes et pour chacun des ces échantillons, nous avons soustrait la valeur de son expression de celle de la moyenne.

Tableau 2.3 Gènes sélectionnés pour valider le résultat des micropuces à ADN

	Gènes	Nombre de couples d'amorces pour un même gène	Nombre de transcrits représentant le gène (Ensembl)	Positionnement de la séquence Affymétrie par rapport aux transcrits
Spécifiques aux leucémies T	<i>Mettl1</i>	1	1	Englobe plusieurs exons
	<i>NLN</i>	1	2	- 1 ^{er} transcrit : répartie sur trois exons ; - 2 ^{ème} transcrit : une partie englobe deux exons et une partie se trouve au niveau du UTR du côté 3'
Spécifiques aux leucémies B	<i>Antl2</i>	2	1	Au niveau du UTR du côté 3'
	<i>Gmp6a</i>	2	1	Au niveau du UTR du côté 3'
Sur-exprimés dans les leucémies T et B	<i>Neil3</i>	2	1	entre deux exons
	<i>Celsr1</i>	1	1	Au niveau du UTR du côté 5'
Sous-exprimés dans les leucémies T et B	<i>Dirc2</i>	2	1	Englobe une partie du dernier exons et une partie du UTR du côté 3'

Comme le montre ce tableau, nous avons choisi deux gènes spécifiques aux leucémies T et deux autres spécifiques aux leucémies B. L'ensemble de ces 4 gènes se trouve sur- exprimé

dans les leucémies par rapport à leurs contrôles. Nous avons aussi sélectionné des gènes communs aux leucémies T et B et qui sont soit sur-exprimés (2 gènes) soit sous- exprimés (1 gène) par rapport aux contrôles T et B. Au niveau de la 3^{ème} colonne, nous avons indiqué le nombre de paires d'amorces utilisées par gène. Par ailleurs et au niveau de la 4^{ème} colonne, nous avons déterminé le nombre de transcrits existant pour chaque gène (disponible dans le site Ensembl). La compagnie Affymétrie met à la disposition des chercheurs la séquence qu'ils ont utilisée pour représenter chaque gène. En comparant cette séquence avec la séquence du transcrit du gène correspondant en utilisant le programme NCBI BLAST, nous pouvons déterminer l'emplacement exact de la séquence Affymétrie par rapport au transcrit (5^{ème} colonne).

Finalement, nous avons testé le niveau d'expression de 7 gènes (*Mettl1*, *Nln*, *Arntl2*, *Gpm6a*, *Neil3*, *Celsr1*, *Dirc2*) par RT-PCR. Afin de déterminer le niveau d'expression de tous ces gènes, nous avons utilisé les mêmes séquences qu'Affymétrie a utilisé pour représenter l'ensemble des gènes présents sur la puce à ADN. Ces séquences sont disponibles sur le site Affymetrix (<https://www.affymetrix.com>). Pour certains gènes (*Arntl2*, *Gpm6a*, *Neil3* et *Dirc2*), nous avons été contraints de commander deux paires d'amorces :

- Pour les gènes *Arntl2* et *Gpm6a*, et d'après la banque de données Ensembl, la séquence Affymetrix se situait au niveau de la région du UTR du côté 3' (voir les figures 2.2 et 2.3). Ainsi, pour le gène *Arntl2*, nous avons choisi une première paire d'amorce avec un oligonucléotide sens situé au niveau de la séquence Affymétrie et un autre anti-sens localisé au niveau du dernier exon du transcrit. La deuxième paire d'amorce était positionnée au niveau de la séquence Affymétrie. Pour le gène *Gpm6a*, nous avons placé la première paire d'amorce au niveau de la séquence Affymétrie alors que la deuxième était située au niveau de deux exons du transcrit de ce gène.

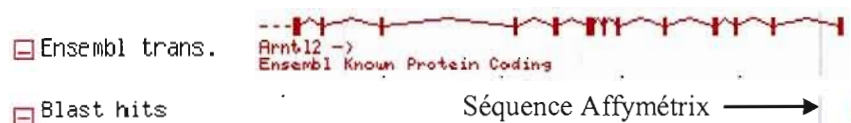


Figure 2.2 Localisation de la séquence Affymétrie au niveau du transcrit du gène *Arntl2*. Ceci a été déterminé grâce au programme Ensembl.



Figure 2.3 Localisation de la séquence Affymétrie au niveau du transcrite du gène *Gpm6a*. Ceci a été déterminé grâce au programme Ensembl.

- Pour le gène *Neil3*, comme le montre la figure 2.4. A, la séquence Affymétrie se situe entre deux exons. Nous avons donc consulté un autre programme appelé UCSC pour voir si notre séquence ne correspondait pas plutôt à un autre transcrite épissé qui n'était pas disponible sur le site Ensembl. Effectivement, un autre transcrite de ce gène correspondait à notre séquence Affymétrie est noté AK031027. Ainsi, nous avons positionné une première paire d'amorce au niveau des deux exons du transcrite disponible sur le site Affymétrie et encadrant la séquence Affymétrie et nous avons positionné une deuxième paire d'amorce au niveau de la séquence Affymétrie elle-même.

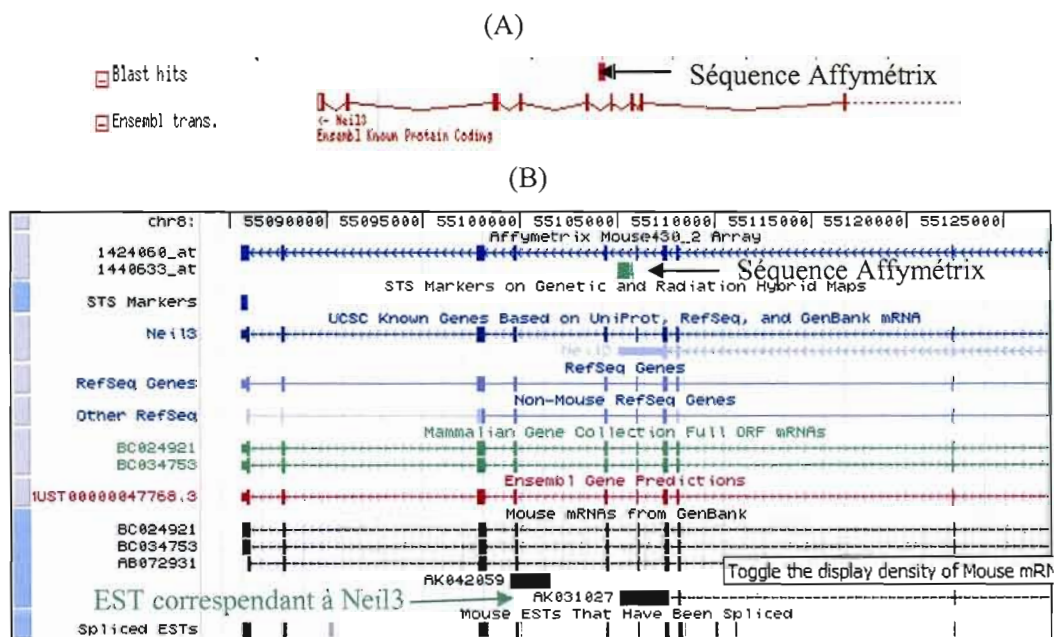


Figure 2.4 Localisation de la séquence Affymétrie au niveau du transcrite du gène *Neil3*. Ceci a été déterminé (A) : grâce au programme Ensembl et (B) : grâce au programme UCSC.

- Une partie de la séquence Affymétrie du gène *Dirc2* se trouve au niveau de la région UTR du côté 3' (*Dirc2* est un gène anti-sens). L'autre partie se trouve au niveau de l'exon en juxtaposition avec le UTR (Voir la figure 2.5). Nous avons donc choisi une première paire d'amorce au niveau des deux derniers exons du transcrit afin d'encadrer la séquence Affymetrie et une seconde au niveau de cette séquence elle-même.



Figure 2.5 Localisation de la séquence Affymetrie au niveau du transcrit du gène *Dirc2*. Ceci a été déterminé grâce au programme Ensembl.

Dans le tableau 2.3 se trouve l'ensemble des 7 gènes sélectionnés pour confirmer le résultat des micropuces à ADN.

2.10 La Reverse Transcriptase-PCR (RT-PCR)

L'ARNm est d'abord extrait à partir des organes congelés selon la méthode du Trizol déjà décrite, ensuite purifié (Kit RNeasy®, QIAGEN) puis rétro-transcrit en ADNc selon le protocole Omniscript (QIAGEN).

Brièvement, pour un volume final de 20 µl, 50 ng à 2 µg d'ARNm sont rétro-transcrits avec 4 unités de l'enzyme Reverse-Transcriptase dans du tampon spécifique à cette enzyme (offert par le Kit de la reverse transcription) en présence de 50 mM de chacun des dNTP (10 mM) et de 1 µM d'un Oligo-dT comme amorce. Le tout est incubé pendant 1 heure à 37 °C. L'inactivation de la Reverse Transcriptase se fait pendant 5 min à 93 °C.

L'expression de chacun des gènes candidats est vérifiée par PCR selon le protocole Omniscript. Ainsi l'ADNc obtenu par rétrotranscription est amplifié en présence d'amorces spécifiques choisies pour chacun des gènes à l'aide d'une Taq polymérase (2.5 Unités) et de dNTP. Parallèlement, un contrôle interne représenté par l'actine («House Keeping gene») est amplifié aux fins de normalisation. Pour exclure toute amplification résultant d'une contamination de l'ARN par de l'ADN génomique, un contrôle négatif sans produit de la

Reverse Transcriptase (c'est à dire sans ADNc) a été réalisé. Les amorces faisant entre 20 et 30 pb ainsi que le cycle thermal de la PCR sont spécifiques à chacun des gènes que nous allons amplifier. Les produits de PCR obtenus sont séparés sur des gels de type non dénaturants.

2.10.1 Quantification du niveau d'expression des gènes GATA-1, GATA-2 et GATA-3 par Reverse Transcriptase-PCR semi-quantitative

Afin d'étudier le niveau d'expression des gènes *GATA-1*, *GATA-2* et de *GATA-3*, l'ARN total de la rate a été extrait. Chacun de ces gènes est spécifique d'une lignée hématopoïétique donnée. Une RT-PCR semi-quantitative a été réalisée grâce aux oligonucléotides spécifiques à chacun de ces gènes (*GATA-1* : 5'-CCAGTGTTCCCATGGATTTT-3' (sens), 5'-TAGAGTGCCGTCTTGCCATA-3' (antisens); *GATA-2* : 5'-TAACAGGCCACTGACCATGA-3' (sens), 5'-ACCACCCT-TGATGTCCATGT-3' (antisens) et *GATA-3* : 5'- CCAAGAGCAGCTCCTTCAAC-3' (sens), 5'- TTCAGTGGTTGGAATGCAGA-3' (antisens). Comme contrôle interne, nous avons choisi l'actine dont les amorces sont : 5'-TGACGGGGTCACCCACACTG-3' (sens), 5'-CTAGAAGCATTTGCGGTGCA-3'.

Pour chaque paire d'amorce, la mise au point de la RT-PCR semi-quantitative a été réalisée en testant différentes quantités d'ARN, différentes températures et différents cycles.

Les cycles de PCR sont constitués de :

- 94°C pendant 3 minutes;
- 94°C pendant 45 secondes;
- Température d'appariement variable selon l'amorce utilisée, pendant 45 secondes;
- 72°C pendant 30 secondes;
- 72°C pendant 10 minutes (pour l'actine : 72°C pendant 3 minutes).

Les températures d'appariement utilisées pour chaque paire d'amorces ainsi que les cycles sont présentés dans le tableau 2.4 :

Tableau 2.4 Différentes températures et différents cycles utilisés pour chaque paire d'amorce

	TM (°C)	Nombre de cycles
<i>GATA-1</i>	55.1	27
<i>GATA-2</i>	61.1	25
<i>GATA-3</i>	56.2	24
Actine	56.7	23

La quantification de l'expression des nos gènes a été réalisée grâce au programme Quantity One. Ce dernier permet de déterminer la densité de chaque bande. Le ratio de densité que nous avons obtenu a été calculé à l'aide de la formule suivante :

$$(\text{Densité d'un gène} - \text{Densité du background}) / (\text{Densité de l'actine} - \text{Densité du background})$$

Ce ratio nous a permis de normaliser nos valeurs et de comparer l'expression de chaque gène entre différents types de leucémies.

2.10.2 Confirmation du résultat des micropuces par Reverse Transcriptase-PCR

Cette RT-PCR a été réalisée avec des amorces spécifiques de chacun des gènes sélectionnées. Ces amorces sont présentées au niveau du tableau 2.5 :

Tableau 2.5 Séquence des paires d'amorces utilisées pour chacun des 7 gènes testés par RT-PCR

Gènes	Paire d'amorces	
	Sens	Antisens
Mettl1	5'-TCGGACTATGTGCAGGACAG-3'	5'-CACATCCACTCATGCAGCTC-3'
Nln	5'- TTGACCAGTCCCTCCATACC-3'	5'- CCCAGGCTTCAGGATTAGGT-3'
Arntl2.1	5'-CCTGGGGACTTC-AGTGACAT-3'	5'-CTGTCACCTGAGCTGCATTC-3'
Arntl2.2	5'-CCCTAAAATCATGAGGCAATG-3'	5'-CTGTCACCTGAGCTGCATTC-3'
Gmpa6.1	5'-TGGACCAAATTCGTGTTTG-3'	5'- CATGCCACAGAAACGCTAGT-3'
Gmpa6.2	5'-GTTCAATTGTGGCACTTGCTG-3'	5'-CGCTCTTTGGAGCGAGTAGA-3'
Neil3.1	5'-GGTGGAA-AGCCAACAGAGAG-3'	5'-GAATGCTGAAATCCCGAGTC-3'
Neil3.2	5'-GACCTCCGTAAAAGTGCTTCA-3'	5'-ATGCCCAATTTGATGTAGGC-3'
Celsr1	5'-GTCTGAATCAGGTCCCCAAA-3'	5'-ACTGAAGCTCCCCTGTGCTA-3'
Dir2.1	5'-TCTGTGCTTCAGGGAATCCT-3'	5'-ACATCCTCACGACACTCACG-3'
Dir2.2	5'-ACATGCAATGGGAGAAAAGG-3'	5'-TCAGCTCATCAGTGGGTCAG-3'

Comme contrôle interne, nous avons utilisé l'actine.

Le profil thermal de la PCR est constitué de :

- 94°C pendant 3 minutes;
- 94°C pendant 45 secondes;
- Température d'appariement variable selon l'amorce utilisée, pendant 45 secondes;
- 72°C pendant 30 secondes;
- 72°C pendant 10 minutes (pour l'actine : 72°C pendant 3 minutes).

La température d'appariement et le nombre de cycles que nous avons utilisés pour l'ensemble des gènes sélectionnés étaient de 56 °C et de 30 cycles respectivement, alors que pour l'actine, ils étaient de 56.7 °C et de 23 cycles.

CHAPITRE III

RÉSULTATS

3.1 Caractérisation des leucémies

Plusieurs critères ont été utilisés pour caractériser les différents types de leucémies observées : signes cliniques et morphologie des tissus, immunophénotypage par l'appareil du FACSvantage et analyse moléculaire.

3.1.1 Analyse des signes cliniques chez la souris et de la morphologie des organes

Tel que décrit par Ru et collaborateurs (1993), nous avons pu confirmer que les souris inoculées avec l'un ou l'autre des deux variants rétroviraux du rétrovirus murin Graffi (GV-1.2 ou GV-1.4) développent la leucémie après une période de latence plus longue pour GV-1.4. Effectivement et malgré le fait que nous n'avons analysé qu'une souche de souris inoculée avec GV-1.2, le tableau 3.1 montre que la souche Balb/c injectée avec GV-1.2 présente un temps de latence plus court (souris Cc64.1: 47 jours) par rapport à toutes celles injectées avec GV-1.4 (souris Cc60.1 : 83 jours, souris Cc60.5 : 120 jours et souris Cc77.3 : 187 jours). Nous avons observé que, pour ces souris, les signes cliniques apparaissent d'une façon subite et non progressive. Cela veut dire que les signes commencent à être observés juste quelques jours avant qu'elles n'atteignent le stade terminal de la maladie. Selon que la leucémie est lymphoïde ou non lymphoïde, les signes cliniques du premier type sont représentés par l'apparition de gros ganglions ou de thymome se traduisant par une respiration lente et difficile. Le second type de signes est souvent caractérisé par l'observation d'une anémie sévère révélée par l'hématocrite s'accompagnant de l'apparition d'oreilles blanches. Comme le montre le tableau 3.1 (colonne 7), la morphologie des organes

internes de ces souris est très hétérogène. Généralement, que ce soit pour les leucémies lymphoïdes ou non lymphoïdes, la taille de rate et du foie se développent beaucoup plus qu'une rate ou qu'un foie d'une souris normale et prennent une couleur variant entre le rouge pâle, le rouge et le rouge foncé. Quelques fois, des tâches blanches peuvent être observées au niveau de ces deux organes. Dans le cas d'une leucémie lymphoïde de type T nous pouvons observer que le développement d'un thymome est souvent accompagné de l'apparition de ganglions lymphatiques, tandis qu'en cas d'une leucémie lymphoïde de type B, l'apparition des ganglions ne s'accompagne pas de l'apparition de thymome. L'hématocrite des souris ayant développé une leucémie de type T est beaucoup plus élevé (entre 32 et 60%) que celui des souris ayant un autre type de leucémie spécifiquement lorsqu'il s'agit d'une leucémie érythroïde (entre 10% et 30%) ou mégakaryocytaire (moins que 15%). Parfois, lorsqu'il s'agit d'une leucémie biphénotypique ou bien d'une leucémie mixte, nous pouvons observer certaines anomalies soit au niveau du cerveau (qui peut devenir soit bombé (souris Cc57.8) soit sanglant (souris Cc58.3)) soit au niveau des reins (qui prennent une couleur pâle (souris Cc57.3) ou en augmentant de taille (souris Cc57.5)).

3.1.2 Analyse des cellules leucémiques par cytométrie de flux

Afin de caractériser le type de leucémie, nous avons immunophénotypé, par cytométrie de flux, trois organes hématopoïétiques pour chacune des souris analysées. La rate et la moelle osseuse sont considérées comme étant des organes essentiels à analyser puisqu'une splénomégalie est toujours observée après injection du rétrovirus murin Graffi pour toutes les souris. Il en va de même pour le sang. Il est à noter qu'il est possible de remplacer ce dernier par le thymus ou par les ganglions selon leur présence ou non et selon l'importance de leur taille. Dans le cadre de ce travail, nous nous sommes limités à l'analyse d'une trentaine de souris. Comme le montre le tableau 3.1 (colonne 8) ci-après, les phénotypes que nous avons obtenus étaient très diversifiés dépendamment de la souche de souris utilisées. Ce tableau témoigne de la complexité et de la diversité des leucémies induites par le rétrovirus murin Graffi. Ces leucémies peuvent être lymphoïdes ou non lymphoïdes. Bien que le nombre d'échantillons ne soit pas très élevé pour chaque type de souris injectée avec GV-1.4, nous avons pu déterminer la proportion de chaque type de

leucémie. La souche Balb/c donne un pourcentage élevé de leucémies lymphoïdes (75% de type T) et une faible part de leucémies bi-phénotypiques (25%). La souche FVB/n donne un pourcentage élevé de leucémies non lymphoïdes (80% de leucémies érythroïdes) et un faible pourcentage de leucémies mixtes (20%), alors que dans la souche NFS injectée avec GV-1.4, nous avons observé la prédominance de leucémies mixtes et bi-phénotypiques formées de leucémies E(Mk), E(B), BT, T(M), M(Mk), M(c-Kit), BMk. Le reste était formé à concurrence de 23% de leucémies de type T, d'environ 7.7% de chacune des leucémies érythroïdes (E) et myéloïdes (M). Deux souris parmi les 30 avaient développé des pathologies autres que les leucémies se traduisant par une paralysie des membres postérieurs et d'un dysfonctionnement du système urinaire. Des souris non infectées par le rétrovirus murin Graffi (5 souris notées par « Cc » pour contrôle) ont aussi été analysées pour chacune des souches de souris et ont servi comme souris contrôles.

Tableau 3.1 Liste des souris étudiées et des caractéristiques des tumeurs induites par le virus Graffi

Souche de souris	N ^o de Souris	Injectées par		Sexe	Période de latence (jours)	Signes cliniques et morphologie des organes internes	Type de leucémies
		GV 1.2	GV 1.4				
Balb/c	Cc60.1		x	F	83	R : (2+) rouge G : (2+) partout Thy : (3+) F : (2+) gonflé S : 60%	T : CD3, CD90, CD4 low (22%) (triée avec CD4)
	Cc60.5		x	F	120	R : (4+) très foncée G : (2+) partout Thy : (1.5+) assez petit F : (2+) rouge S : 45%	Biphénotypique : TM : CD3, CD90, CD4, Mac1 low (triée avec CD4)
	Cc64.1	x		F	47	R : (3+) rouge foncé G : (2+) partout Thy : (3+) gros F : (1+) normal S : 45%	T : CD3, CD90, CD4
	Cc77.3		x	F	187	R : (5+) rouge G : (1+) sur les côtés Thy : (3+) F=gros	T : CD3, CD90, CD4, Mac1 low TM (triée avec CD4)
	Cc77.4		x	M		Autres problèmes Grosse vessie	Non analysée
	Contrôle Cc02	-	-	M	-	-	-
	Contrôle Cc03	-	-		-	-	-
	Cc55.2		x	F	118	R : (2+) rouge F : (3+) énorme S : 10%	E : Ter119, CD71, CD41 low TM

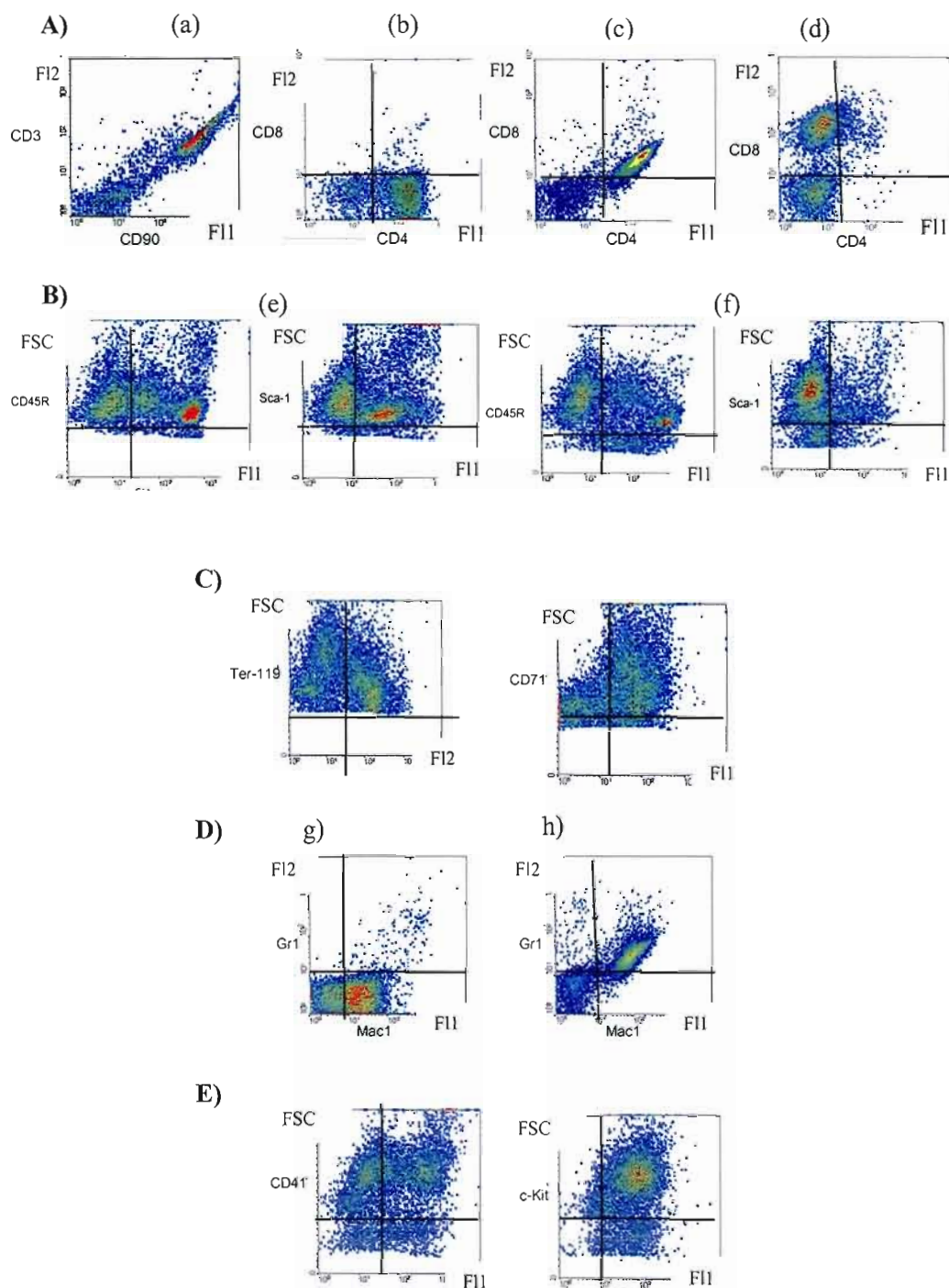
FVB/n	Cc59.1		x	F	104	R : (5+) G : (3+) partout F : (4+) énorme S : 30%	E (non invasive) Ter119, CD71, Scal, c-kit, CD41 low
	Cc59.2		x	F	99	R : (3+) rouge F : (4+) granuleux S : 20%	E Ter119, CD71, c-kit (triée avec TER- 119)
	Cc59.5		x	M	92	R : (3+) rouge F : (1+) normal Reins : pâles S : 5%	Mixte BMk : CD45R, Mk, c-Kit (Triée avec CD 41)
	Cc62.2		x	F	43	R : (2+) assez pâle F : (3+) granuleux	E (non invasive) Ter119, CD71, Scal
	Contrôle Cc 01	-	-	-	-	-	-
NFS	Cc57.1		x		105	R : (4+) pâle, granuleuse G : (3+) partout Thy : (3+) F : (4+) infiltré	T: CD3, CD90, CD4
	Cc57.2		x		105	R : (4+) blanche, granuleuse G : (2+) partout Thy : (3+) Foie : (4+) infiltré S : 35%	T: CD3, CD90, CD4 low
	Cc57.3		x	M	112	R : (3+) rouge foncé F : (3+) infiltré Reins : pâles S : 15%	Biphénotypique E(Mk) : Ter119, CD71, c-kit low, scal low

Cc57.5		x	M	137 jours	R : (3+) rouge G : (2+) partout Thy : (2+) Foie : (1.5+) normal Riens : (1.5) S : 74%	biphénotypique : T(M) : CD3, CD90, CD4, Mac1 (triée avec Mac-1)
Cc57.6		x	M	155	R : (3+) rouge G : (3+) surtout au niveau des intestins Thy : (2+) F : (4+) pâle Riens : (1.5) S : 20%	Mixte E(B) : - Rate : Ter119, CD71, c-kit, Gr1 low - Ganglions : CD45R, Scal, c-kit
Cc57.8		x	F	168	R : (3+) pâle F : (2+) non infiltré avec des tâches blanches Reins : très pâles C : bombé S : 15%	Mixte BMk : - Rate : M, Mk, Mac1, Gr1, CD41 - Moelle osseuse : Mac1, Gr1, c-kit, CD41 - Sang : Mac1, Gr1, c-kit, CD41 (triée avec Mac-1)
Cc58.1		x	M	143	R : (3+) rouge foncé G : (1.5+) pas nombreux, dispersés Thy : (2+) F : rouge foncé S : 32%	T : CD3, CD90, CD4, Mac1 low (triée avec CD4)
Cc58.3		x	M	152	R : (2+) moyennement rouge Thy : (2+) F : normal C : sanglant S : 65%	- Rate+moelle osseuse : B - Thymus : T (pas au stade terminal ou pas invasive)
Cc58.5		x	M	179	R : (3+) rouge foncé F : (1.5+) pâle S : 8%	Mixte MMk : Mac1, c-kit, CD41 (triée avec Mac-1)

	Cc58.6		x	F	187	R : (4+) foncée G : (1+) sur les côtés F : (1+) pâle Riens : (2+) très pâles S : 7%	MMk mixte: Mac1, Gr1, CD41, ckit
	Cc61.1		x	F	111	R : (2+) pas très infiltré G : (3+) partout F : (1+) pas très infiltré S : 50%	Mixte B+ Még (Rate et Thymus) (triée avec CD45R)
	Cc61.2		x	F	163	R : (4+) infiltré F : (2+) pas très infiltré Riens : pâles S : 10%	Mixte MMk(B) : Mac1, Gr1, CD41, c-kit, CD45R, Sca1
	Cc76.1		x	M	234	R : (5+) rouge, un peu granuleuse F : (3+) avec des tâches blanches S : liquide	M Mac1, Gr1 (triée par Mac1)
	Cc79.2		x	M		Paralysée grosse vessie	Non analysée
	Contrôle Cc04	-	-	M	-	-	-
	Contrôle Cc05	-	-	F	-	-	-

Ce tableau représentant les différentes souris analysées, les variants rétroviraux injectés, le temps de latence, les signes cliniques et la morphologie des organes ainsi que les résultats obtenus par cytométrie de flux. Abréviations : R, rate ; Thy, Thymus ; G, ganglions ; F, foie ; C, cerveau ; S, sang ; T, leucémie lymphoïde de type T ; B, leucémie lymphoïde de type B ; E, leucémie érythroïde ; M, leucémie myéloïde ; Mk, leucémie mégakaryocytaire. Les leucémies bi-phénotypiques ou mixtes sont représentées par le rassemblement de deux abréviations. Les chiffres entre parenthèses indiquent le nombre de fois d'augmentation de la taille de l'organe en question par rapport à la normale. Uniquement les leucémies non mixtes non bi-phénotypiques ont été colorées : les leucémies T sont en vert ; les E sont en rose ; les M sont en rouge.

La figure suivante illustre quelques exemples de cellules leucémies et normales (contrôles) immunophénotypées par cytométrie de flux :



F)

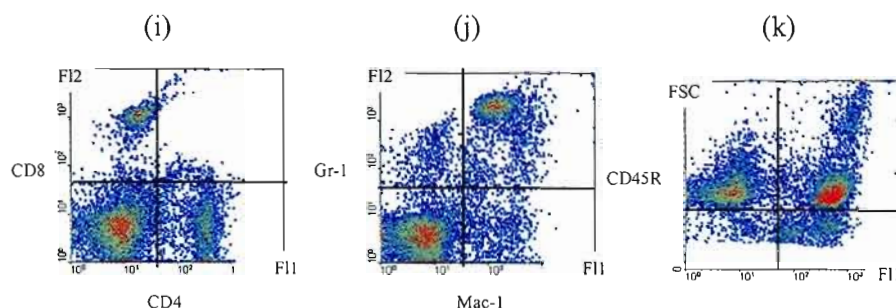


Figure 3.1 Immunophénotypage des cellules leucémiques et normales par cytométrie de flux. Les cellules en suspension sont marquées avec différents anticorps spécifiques à chaque lignée hématopoïétique. Chaque anticorps est marqué avec un fluorophore (FITC (Fluorescéine) ou PE (Phycoérythrine)). Grâce à l'intensité émise par les fluorophores suite à leur excitation, il est possible de déterminer la taille de la population cellulaire marquée avec les anticorps (FSC « forward scatter »), sa granulosité (SSC « side scatter »), sa fluorescence sous divers angles (F11 et F12). Les différentes figures présentées ci-haut, donnent un aperçu sur la diversité des leucémies induites par le rétrovirus murin Graffi. (A) : leucémies lymphoïdes de type T; (a) : $CD3^+CD90^+$; (b) : $CD4^+CD8^-$; (c) : $CD4^+CD8^+$; (d) : $CD4^+CD8^+$. (B) : leucémies lymphoïdes de type B; (e) : $CD45R^+Sca-1^+$; (f) : $CD45R^+Sca-1^-$; (C) : leucémie érythroïde $TER-119+CD71^+$; (D) : leucémies myéloïdes; (g) : $Mac1^+Gr1^+$; (h) : $Mac1^+Gr1^+$; (E) : leucémie mégakaryocytaire $CD41^+c-Kit^+$. (F) : cellules de souris contrôles; (i) : $CD4^+CD8^+$; (j) : $Mac1^+Gr1^+$; (k) : $CD45R^+$. Lorsque les cellules sont doublement marquées (exemple avec les anticorps anti-CD4 et anti-CD8) l'axe des X est représenté par F11 et l'axe des Y est représenté par F12. Par contre, lorsque les cellules sont marquées avec un seul anticorps (exemple avec les anticorps anti-CD45R,) l'axe des X est représenté par F11 et l'axe des Y est représenté par FSC.

3.1.3 Analyse moléculaire

Dans le but de confirmer les résultats obtenus suite à l'analyse par cytométrie de flux, nous avons étudié l'expression des gènes *Fog1*, *GATA-1*, *GATA-2* et de *GATA-3*.

3.1.3.1 Etude de l'expression du gène *Fog-1* par Northern blot

Dans un premier temps, l'expression du gène *Fog-1* a été analysée. Le gène *Fog-1* est considéré comme un marqueur exprimé dans les leucémies E et Mk. Nous avons analysé les ARN de différents types de leucémies provenant de différentes souches de souris (se reporter au tableau 3.1 pour retrouver la souche de souris correspondant aux numéros

indiqués dans la légende de la figure 3.2) après séparation par électrophorèse puis transfert sur membrane de nylon par hybridation avec la sonde *Fog-1* à notre disposition. La figure 3.2 montre que *Fog-1* est très faiblement exprimé au niveau des leucémies lymphoïdes de type T (souris 64.1 et 57.2). Ce niveau d'expression est identique à celui présent au niveau des souris contrôles (souris 2 et 4) et de celles n'ayant pas développé la leucémie (souris 43.5). Par contre, *Fog-1* est très fortement exprimé au niveau de la leucémie mixte E(Mk) (souris 57.3). Pour le restant des leucémies et par rapport aux souris contrôles, un haut niveau d'expression de *Fog-1* est observé. Cependant, ce niveau est deux fois moins important que celui de la leucémie E(Mk) à l'exception de celui indiqué pour la souris 58.6 atteinte d'une forme de leucémie très complexe (MMk(B)). Pour cette souris, le niveau d'expression de *Fog-1* est 4 fois moins important que celui de la souris présentant une leucémie E(Mk).

Remarque

Cette expérience n'a été réalisée qu'une seule fois. Pour cette raison, il nous est impossible de représenter des barres d'erreurs.

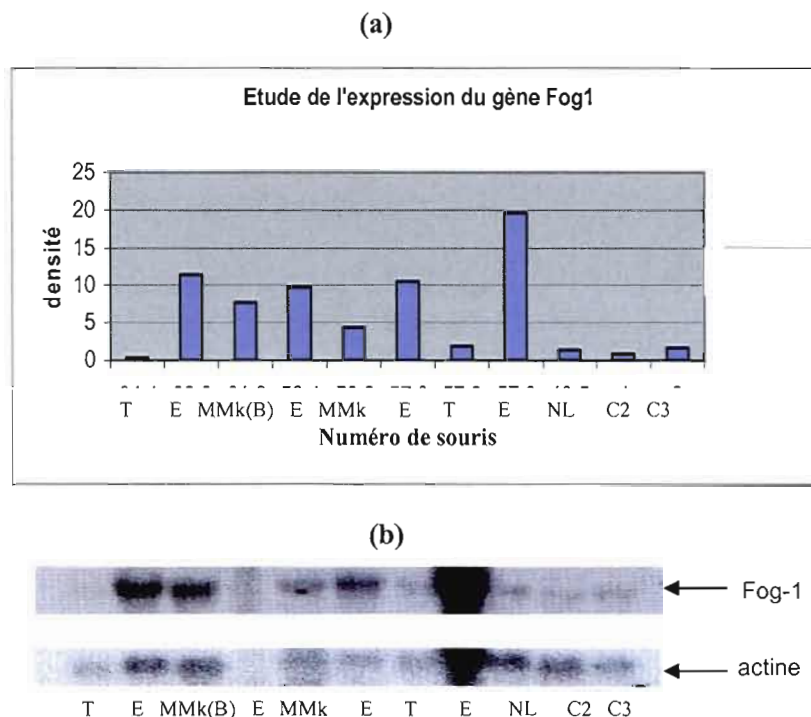


Figure 3.2 Analyse de l'expression du gène Fog-1 par Northern blot pour différents types de leucémies. L'hybridation de type Northern a été faite à l'aide de sonde *Fog-1* (b). Afin de normaliser nos résultats, nous avons utilisé l'actine comme gène de référence et déterminé son niveau d'expression. Ceci nous a permis de construire le graphique présenté dans cette figure (a). 64.1 : leucémie T ; 62.2 : leucémie E (non invasive avec Sca-1) ; 61.2 : leucémie mixte (MMk(B)) ; 59.1 : leucémie E (non invasive) avec Sca-1, CD41 et c-kit ; 58.6 : leucémie mixte MMk ; 57.6 : leucémie E(B) ; 57.2 : leucémie T ; 57.3 : leucémie E(Mk) ; 43.5 : souris n'ayant pas développé de leucémie (NL) ; C2 : souris contrôle 1 ; C3 : souris contrôle 2.

3.1.3.2 Etude du niveau d'expression des gènes GATA-1, GATA-2 et GATA-3 par RT-PCR semi-quantitative

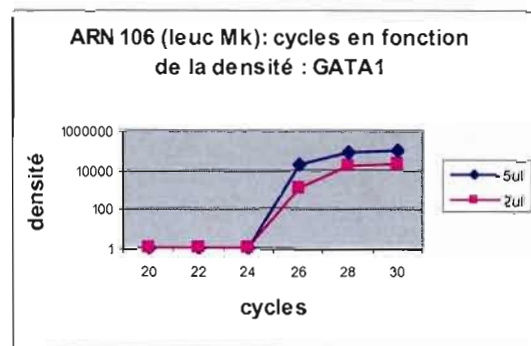
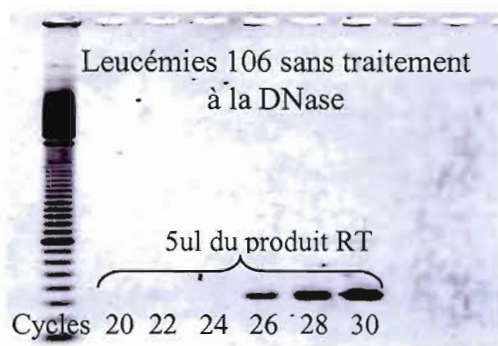
Toujours dans le but de mieux caractériser les cellules leucémiques, nous nous sommes intéressés à l'étude de l'expression de trois gènes *GATA-1*, *GATA-2* et *GATA-3* connus comme étant impliqués dans l'hématopoïèse. Comme contrôle interne, nous avons utilisé l'actine. Pour tester l'expression de ces 3 gènes, nous avons sélectionné 41 différents types de leucémies représentées dans le tableau suivant :

E	EMk	Mk	M	T	B
16	2	11	5	3	4

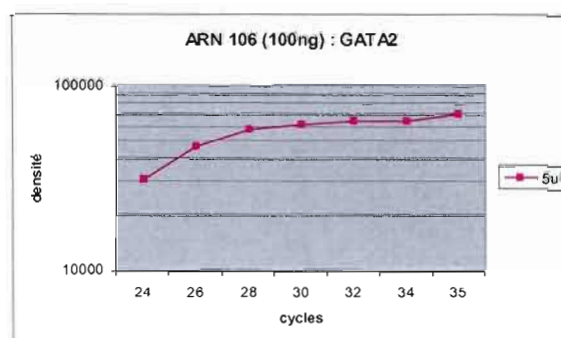
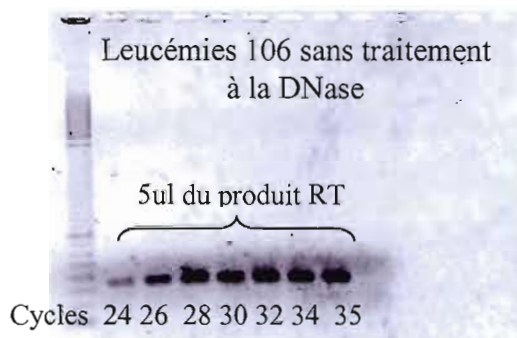
3.3.2.1 Mise au point de la technique de la RT-PCR semi-quantitative pour GATA-1, GATA-2 et GATA-3

La réaction de PCR est une réaction exponentielle. Elle est caractérisée par une phase exponentielle (le nombre de copies d'ADNc augmente d'une façon exponentielle sous la forme de 2^n) suivie d'un plateau (lorsqu'il n'y a plus de production de copies d'ADNc : c'est la saturation). Afin de faire une semi-quantification du niveau d'expression de nos 3 gènes, il fallait que le nombre de cycle choisis nous permette d'obtenir une quantité d'ADNc se situant dans la phase exponentielle. Comme le montre la figure 3.3. A, les copies d'ADNc sont détectées à partir du 24^{ème} cycle alors que le plateau est atteint à partir du 28^{ème}. Ainsi, pour la semi-quantification du niveau d'expression du gène *GATA-1*, nous avons choisi 27 cycles pour la réaction de PCR. Dans les figures 3.3. B et 3.3. C, nous avons commencé la semi-quantification à partir de du 24^{ème} cycle pour *GATA-2* et à partir du 25^{ème} cycle pour *GATA-3*. Ces deux figures nous ont permis de réaliser que cette semi-quantification aurait dû commencer à partir d'un nombre de cycles inférieurs à celui que nous avons choisi pour chaque gène étant donnée que la phase exponentielle n'est pas évidente comme celle pour *GATA-1* (la détection de copies d'ADNc est beaucoup plus précoce que celle de *GATA-1*). Cependant, nous pouvons voir que pour le gène *GATA-2*, la phase exponentielle est atteinte à partir du 24^{ème} cycle (figure 3.3. B) alors qu'elle est atteinte pour le gène *GATA-3* qu'au 25^{ème}. Le plateau débute environ au 30^{ème} cycle pour ces deux gènes et ce malgré que nous pouvons noter une diminution du nombre de copies d'ADNc pour le gène *GATA-3* (cela est peut être expliquée par une imprécision de la manipulation). Ainsi, pour amplifier les gènes *GATA-2* et *GATA-3*, nous avons choisi 25 cycles pour le premier et 24 cycles pour le second.

(A) - Mise au point de la technique de la RT-PCR pour le gène GATA-1.



(B)- Mise au point de la technique de la RT-PCR pour le gène GATA-2.



(C)- Mise au point de la technique de la RT-PCR pour le gène GATA-3.

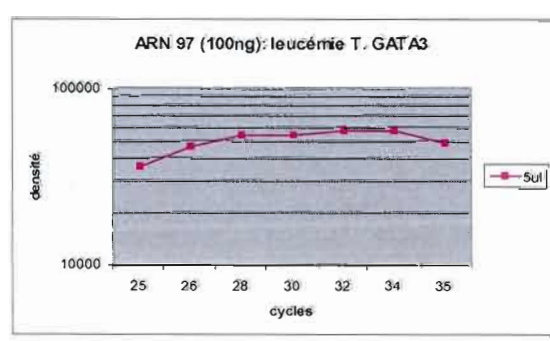
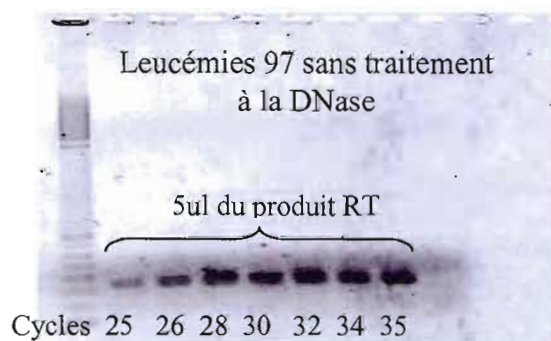


Figure 3.3 Optimisation de la réaction de PCR pour les gènes GATA-1, GATA-2 et GATA-3. Pour que notre réaction de PCR soit semi-quantitative, nous avons testé différents cycles pour le gène *GATA-1* (figure A), *GATA-2* (figure B) et *GATA-3* (figure C).

3.1.3.2.2 Etude du niveau d'expression des gènes GATA-1, GATA-2 et GATA-3 au niveau de différents types de leucémies par RT-PCR semi-quantitative

Suite à l'amplification du gène *GATA-1* par RT-PCR semi-quantitative et d'après la figure 3.4.A, nous avons observé que *GATA-1* semble être très fortement exprimé dans l'ensemble des échantillons excepté dans les leucémies Mk et EMK. La figure 3.4.B montre qu'au niveau des leucémies lymphoïdes de type T, ce gène se trouve le plus fortement exprimé. De même, *GATA-1* se trouve fortement exprimé dans les leucémies myéloïdes mais à un niveau moins élevé que celui des leucémies T. Les écarts-types montrent qu'il existe une variabilité entre les différents échantillons d'un même type. Cependant, cette dernière est assez importante spécifiquement pour les leucémies lymphoïdes de type B.

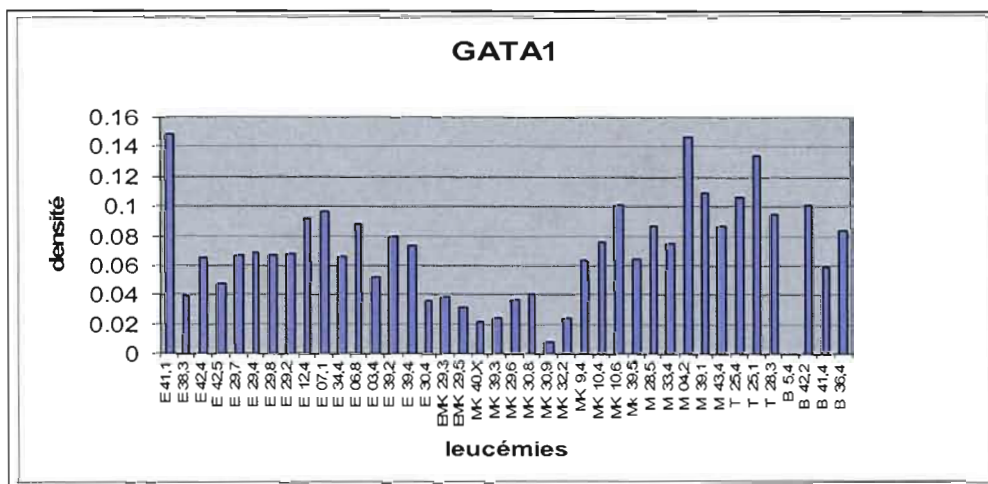
Les figures 3.5.A et 3.5.B montrent que le gène *GATA-2* semble plus exprimé dans les leucémies non lymphoïdes que dans les leucémies lymphoïdes. Les écarts-types montrent que pour les leucémies EMk et M, la variance est très importante.

Dans la figure 3.6.A, le niveau d'expression du gène *GATA-3* semble moins homogène comme le montrent aussi les écarts-types de la figure 3.6. B (cette variance semble toutefois moins importante pour les leucémies T). Cependant, au niveau de cette dernière, nous avons remarqué que ce gène est le plus fortement exprimé au niveau des leucémies T surtout par rapport aux leucémies EMk où il ne s'exprime pas. Le niveau d'expression de *GATA-3* au niveau des leucémies B représente un peu moins que 50% par rapport aux leucémies T. Par contre l'expression de *GATA-3* chez les M et des E est respectivement 5 à 12 fois moins importantes que les leucémies T.

Remarque

L'apparition de tels écarts-types n'est pas étonnante étant donné que les leucémies forment un groupe très hétérogène. Par exemple, pour les leucémies lymphoïdes de type B, certaines peuvent exprimer le marqueur CD19 uniquement, d'autres peuvent exprimer le marqueur Sca-1 en plus du CD19. D'autres peuvent exprimer CD71 en plus de ces deux marqueurs. Ceci ne représente qu'une petite démonstration des différentes combinaisons qui peuvent exister pour les cellules lymphoïdes de type B.

(A)



(B)

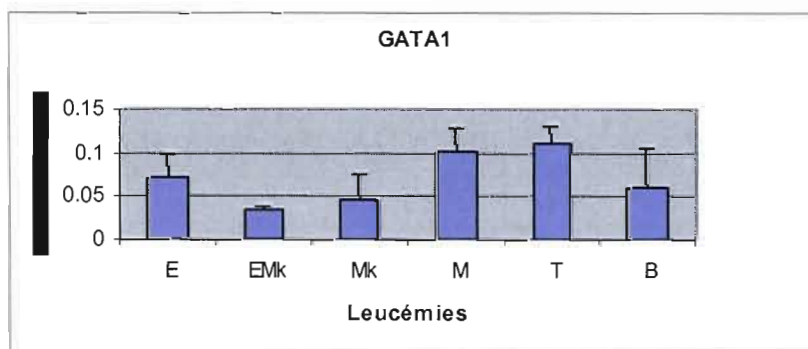
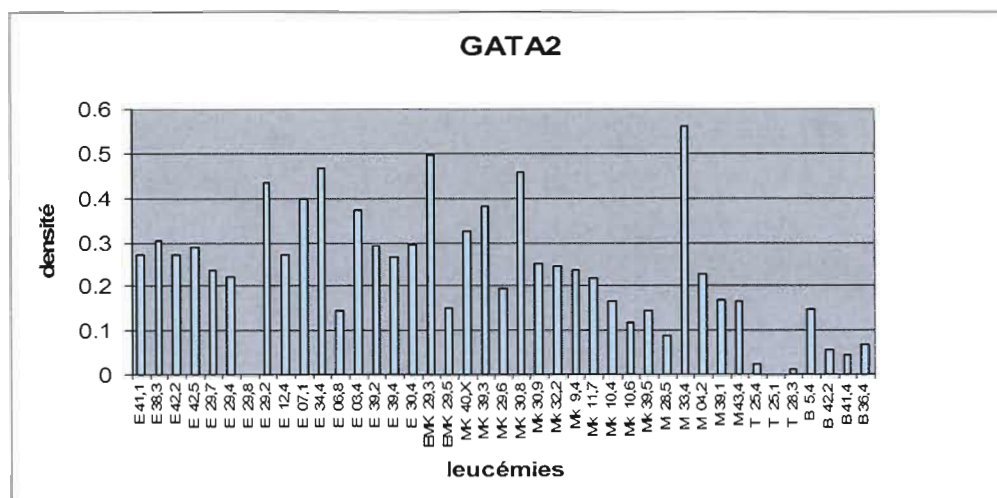


Figure 3.4 Etude du niveau d'expression du gène GATA-1 au niveau de différents types de leucémies. Nous avons étudié le niveau d'expression du gène *GATA-1* au niveau de 41 échantillons (16 leucémies érythroides (E), 2 leucémies érythroides mégakaryocytaires (EMk), 11 leucémies mégakaryocytaires (Mk), 5 leucémies myéloïdes (M), 3 leucémies T et 4 leucémies B). Afin de normaliser nos résultats, nous avons utilisé l'actine comme gène de référence et déterminé son niveau d'expression. Ceci nous a permis de construire l'ensemble des graphiques présentés dans cette figure. Par ailleurs et pour avoir une vue globale, nous avons calculé la moyenne d'expression de chacun de ces gènes au niveau de toutes les leucémies faisant partie d'un même type (E, EMk, Mk, M, T, B). (A) : Niveau d'expression du gène *GATA-1* au niveau des 41 leucémies; (B) : Moyenne de l'expression de *GATA-1* au niveau des 41 leucémies étudiées pour chaque groupe de leucémies.

(A)



(B)

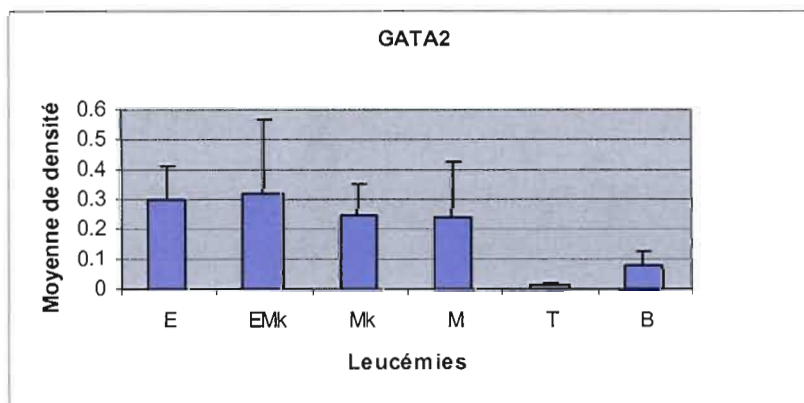
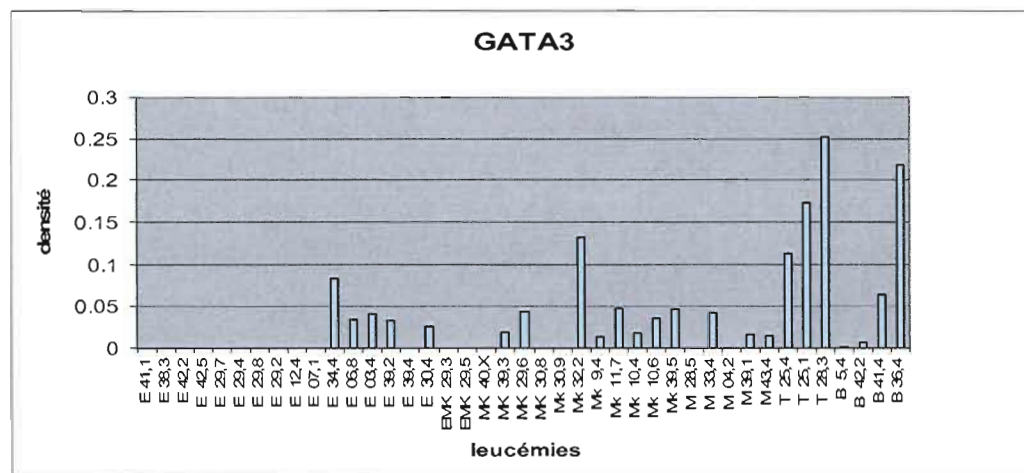


Figure 3.5 Etude du niveau d'expression du gène *GATA-2* au niveau de différents types de leucémies. Pour étudier le niveau d'expression du gène *GATA-2* au niveau de 41 leucémies, nous avons procédé de la même façon décrite pour le gène *GATA-1*. (A) : Niveau d'expression du gène *GATA-2* au niveau des 41 leucémies; (B) : Moyenne de l'expression de *GATA-2* au niveau des 41 leucémies étudiées pour chaque groupe de leucémie.

(A)



(B)

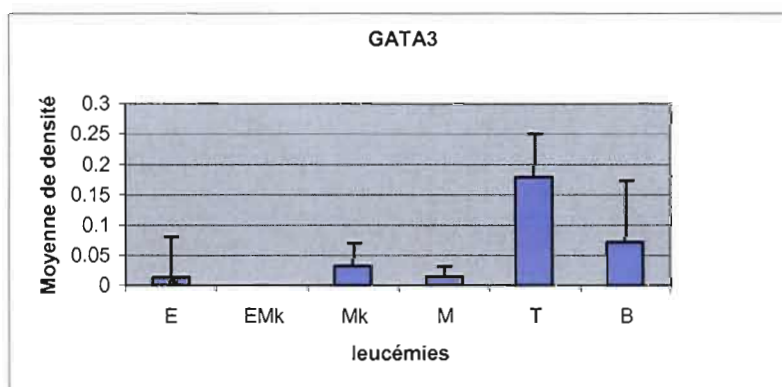


Figure 3.6 Etude du niveau d'expression du gène *GATA-3* au niveau de différents types de leucémies. Pour étudier le niveau d'expression du gène *GATA-3* au niveau de 41 leucémies, nous avons procédé de la même façon décrite pour le gène *GATA-1*. (A) : Niveau d'expression du gène *GATA-3* au niveau des 41 leucémies; (B) : Moyenne de l'expression de *GATA-3* au niveau des 41 leucémies étudiées pour chaque groupe de leucémie.

3.2 Analyse des différentes leucémies induites par le rétrovirus Graffi à l'aide des micropuces à ADN

3.2.1 Tri des cellules leucémiques selon leurs caractéristiques phénotypiques

Le tri des cellules a été effectué à l'aide de billes magnétiques par une sélection positive avec les anticorps spécifiques de la lignée leucémique à purifier. La figure 3.7.A représente un exemple de tri des cellules d'une leucémie T ($CD4^+/CD8^-$). Suite aux étapes de lavage nous permettant de nous débarrasser des populations non désirées, le surnageant contenant ces dernières est analysé par cytométrie de flux (figure 3.7.B). Ce surnageant est constitué de cellules normales, de cellules mortes et de cellules provenant d'autres populations leucémiques. Lors des lavages, un faible pourcentage des cellules appartenant à la population à purifier, est éliminé. La figure 3.7.C représente la sélection positive où environ 90% des cellules récupérées sont représentés par la population à purifier et environ 10% sont représentés par d'autres types cellulaires (un faible pourcentage de contamination).

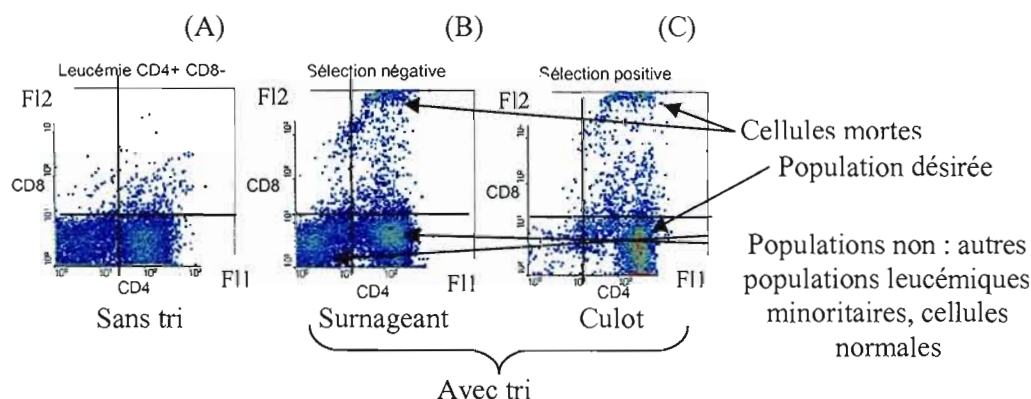


Figure 3.7 Exemple d'un tri d'une leucémie $CD4^+/CD8^-$ par la technique des billes magnétiques. (A) : Immunophénotypage du type de leucémie par cytométrie de flux. Les cellules ont été triées par incubation avec des billes magnétiques sur lesquelles sont fixés les anticorps spécifiques puis mis au contact d'un aimant ; (B) : immunophénotypage des cellules récupérées dans le surnageant (sélection négative) ; (C) : immunophénotypage de cellules dans le culot au contact de l'aimant (sélection positive).

Le tableau 3.2 présente les différentes leucémies triées (17 leucémies) par la technique des billes magnétiques. Par ailleurs, nous avons inclus dans le tableau 3.2 les 6 leucémies que nous avons envoyées au centre génomique de Montréal pour être hybridées sur les micropuces à ADN. Pour ces derniers, nous avons trié un échantillon contrôle pour les leucémies B $CD45R^+/CD19^+$ et un autre pour les leucémies T $CD4^+ CD8^+$. Pour les leucémies T et B, nous avons trié différents sous types. C'est ainsi par exemple qu'au niveau des leucémies T nous avons trié des leucémies $CD4^+/CD8^-$, $CD8^- CD4^+$ et $CD4^+/CD8^+$.

Tableau 3.2 Ensemble des leucémies triées par la technique des billes magnétiques

Type des leucémies triées	Triées avec les anticorps anti-
6 leucémies T	CD4 et CD8
4 leucémies B	CD45R
1 leucémie E (érythroïde)	Ter119
1 leucémie M (myéloïde)	Mac-1
2 leucémies biphénotypiques BMk	CD45R
2 leucémies biphénotypiques TM	CD4
1 leucémie mixte MMk	Mac-1

La première colonne représente tous les échantillons triés et la deuxième contient les anticorps ainsi que les combinaisons utilisées pour trier différents sous types de leucémies T ou B.

Remarque

L'anticorps CD45R utilisé est exclusivement spécifique aux leucémies lymphoïdes de type B. Il est important de ne pas le confondre avec l'anticorps CD45 qui en plus d'être exprimé au niveau des cellules B, est aussi exprimé au niveau des cellules T.

3.2.2 Micropuces à ADN

L'ARN des cellules leucémiques et des cellules contrôles a été utilisé comme sonde sur des micropuces à ADN Affymetrix dont le «set» est nommé « The Genechip® Mouse Genome 430 2.0 Array set ». Ceci nous a permis d'avoir le profil d'expression de plus de 39 000 transcrits. Afin de sélectionner des gènes qui seraient plus spécifiques aux leucémies T et B ou communs à ces deux types de leucémies, nous avons appliqué différents niveaux de sélection.

Pour mieux étudier et caractériser les leucémies lymphoïdes de type T et B, la stratégie que nous avons adoptée consistait à déterminer :

- les gènes qui sont différentiellement exprimés entre les leucémies T par rapport aux cellules T contrôle mais qui ne varient pas lors de la comparaison entre les leucémies B et des cellules B contrôle. Ces gènes représentent ceux qui sont spécifiques aux leucémies T;
- les gènes qui sont différentiellement exprimés entre les leucémies B par rapport aux cellules B contrôles mais qui restent inchangés entre les leucémies T et leur contrôle. Ces gènes sont considérés comme étant spécifiques aux leucémies B;
- les gènes qui sont différentiellement exprimés par rapport aux contrôles T et B et communs entre les leucémies T et B. Ces derniers peuvent être divisés en deux sous-groupes selon qu'ils soient sur-exprimés ou sous-exprimés dans les deux types de leucémies par rapport à leur contrôle respectif.

3.2.2.1 Première sélection des gènes d'intérêt

D'abord, nous avons comparé les leucémies lymphoïdes avec leur contrôle respectif. Les gènes que nous avons choisis devaient varier d'un minimum de 4 fois entre les leucémies par rapport à leur contrôle respectif. Ce haut taux de variation entre les leucémies et leur contrôle correspondant « fold-change » (4 fois) était choisi dans le but d'éviter les faux positifs. Ainsi, nous avons trouvé 59 gènes spécifiques aux leucémies T (Annexe C. a), 278

gènes spécifiques aux leucémies B (Annexe C. a) et 212 gènes communs entre les leucémies T et B (Annexe C. b) (pour plus de détails voir le tableau 3.3).

Tableau 3.3 Tableau récapitulatif du nombre de gènes sélectionnés par la méthode RMA

Groupes	Sous-groupes	Nombre de gènes avec un «fold-change» supérieur ou égal à 4
Spécifiques aux leucémies T et ne variant pas dans les leucémies B par rapport au CB	Sur exprimés dans les T par rapport au contrôle T	9
	Sous exprimés dans les T par rapport au contrôle T	50
Spécifiques aux leucémies B et ne variant pas dans les leucémies T par rapport au CT	Sur exprimés dans les B par rapport au contrôle B	190
	Sous exprimés dans les B par rapport au contrôle B	88
Communs entre les leucémies T et les B	Sur exprimés dans les T et les B par rapport à leur contrôle correspondant	94
	Sous exprimés dans les T et les B par rapport à leur contrôle correspondant	118

Le programme RMA nous a permis de classer nos gènes selon qu'ils soient spécifiques ou communs aux leucémies T et B. Ce nombre de gènes qui peut être plus restreint ou plus large, est fonction d'un paramètre nommé «fold-change». Plus le «fold-change» diminue, plus le nombre de gènes, différenciellement exprimés, augmente.

3.2.2.2 Méthode de sélection par «Hierarchical Clustering»

A partir de l'ensemble des gènes qui varient d'un minimum de 4 fois entre les leucémies par rapport à leur contrôle correspondant, nous avons appliqué un programme nommé « Hierarchical Clustering ». Comme le montre la figure 3.8, ce programme nous a

permis de voir que certains gènes varient de la même façon entre les 3 leucémies B et d'autres varient de la même façon entre les 3 leucémies T. Nous pouvons voir que notre algorithme est divisé en 4 parties :

- la première correspond aux gènes qui sont communs aux leucémies T et B et qui sont sous-exprimés par rapport à leur contrôle correspondant;
- la deuxième est constituée des gènes qui sont sur-exprimés dans les 3 leucémies T par rapport au contrôle T;
- la troisième est constituée des gènes qui sont sur-exprimés dans les 3 leucémies B par rapport au contrôle B;
- la quatrième correspond aux gènes qui sont communs aux leucémies T et B et qui sont sur-exprimés par rapport à leur contrôle correspondant.

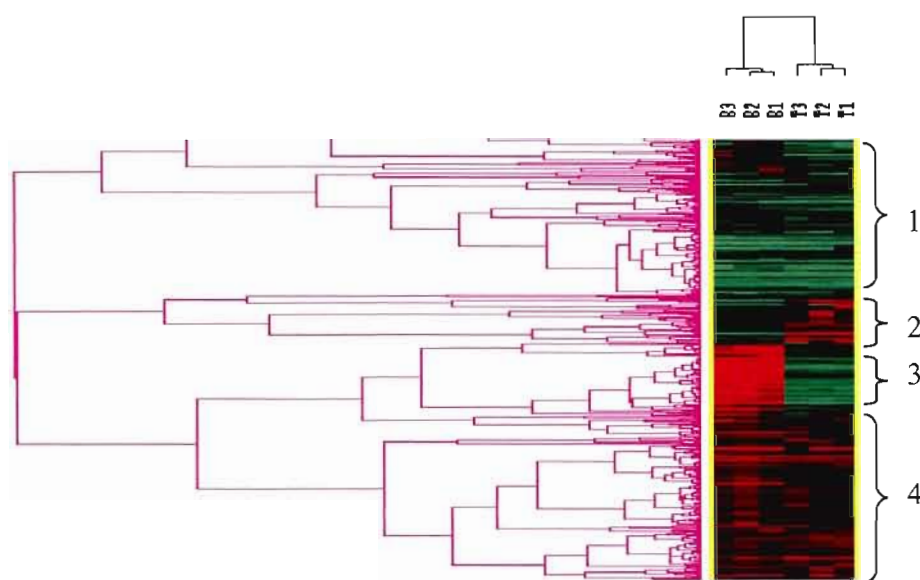


Figure 3.8 Algorithme du « Hierarchical Clustering » des leucémies B et T. Tous les gènes provenant de la première sélection ont été insérés dans le programme du « Hierarchical Clustering ». Ce dernier a permis de les regrouper en fonction de leur degré de similitude en prenant en considération l'intensité de leur expression. Le nom de chaque échantillon est indiqué tout en haut de la figure. Les couleurs correspondent au niveau d'abondance de l'ARN des gènes détectés. Chacun de ces gènes est représenté par une ligne horizontale. **1** : gènes communs entre les leucémies T et B et qui sont sous exprimés par rapport à leurs contrôles correspondants; **2** : gènes qui sont sur exprimés dans les 3 leucémies T par rapport au contrôle T; **3** : gènes qui sont sur exprimés dans les 3 leucémies B par rapport au contrôle B; **4** : gènes qui sont communs entre les leucémies T et B et qui sont sur exprimés par rapport à leur contrôles correspondants.

3.2.2.3 Analyse des gènes sélectionnés

Suite à la première sélection et à partir de l'ensemble des 549 gènes issus de cette dernière, notre recherche bibliographique nous a amené à identifier 56 gènes. Pour ces derniers, notre choix était basé sur le fait qu'ils ne devaient pas avoir été préalablement impliqués dans les leucémies. Ainsi à partir des 549 gènes seuls les gènes qui étaient connus dans différents types de cancers mais pas encore connus au niveau des leucémies ont été sélectionnés, de même ceux qui ont certaines fonctions nécessaires au fonctionnement de l'organisme soit encore qu'aucune fonction ne leur a encore été attribuée. Dans l'annexe D, E et F se trouve l'ensemble de ces 56 gènes dont 31 sont spécifiques aux leucémies B (annexe D), 8 spécifiques aux leucémies T (Annexe E) et 17 communs aux leucémies T et B (annexe F). Ces gènes sont soit sur-exprimés soit sous-exprimés dans les leucémies par rapport à leurs contrôles correspondants.

Notre méthode de sélection nous a aussi permis de retrouver des gènes connus dans la littérature comme étant impliqués que ce soit dans les leucémies de type T ou de type B. Le tableau 3.4 présente quelques exemples de tels gènes :

Tableau 3.4 Niveau d'expression des gènes connus associés à des leucémies T ou B

(A)

ProbesetIDs	B1-CB	B2-CB	B3-CB	Gene Title	Gene Symbol
1419609_at	-3.45	-3.9	-3.755	chemokine (C-C motif) receptor 1	Ccr1
1425289_a_at	-2.819	-3.56	-4.964	complement receptor 2	Cr2
1415849_s_at	2.606	3.023	2.973	stathmin 1	Stmn1
1428142_at	4.04	3.531	3.656	ets variant gene 5	Etv5
1436398_at	4.357	4.485	4.482	Lymphoid enhancer binding factor 1	Lef1
1441242_at	4.739	4.807	4.522	Dipeptidylpeptidase 4	Dpp4
1426817_at	5.641	6.093	5.88	antigen identified by monoclonal antibody Ki 67	Mki67
1423756_s_at	6.57	5.221	6.668	insulin-like growth factor binding protein 4	Igfbp4

(B)

ProbesetIDs	T1-CT	T2-CT	T3-CT	Gene Title	Gene Symbol
1441242_at	2.949	3.767	1.128	Dipeptidylpeptidase 4	Dpp4
1421655_a_at	2.863	1.22	2.479	chemokine (C-C motif) receptor 4	Ccr4
1441552_at	2.621	3.046	1.027	T-cell receptor alpha chain	Tcra
1421965_s_at	2.616	3.477	-2.143	Notch gene homolog 3 (Drosophila)	Notch3
1418634-at	2.312	2.921	-1.294	Notch gene homolog 1 (Drosophila)	Notch1

En faisant une recherche bibliographique, nous avons trouvé que certains gènes ont déjà été identifiés que ce soit au niveau des leucémies lymphoïdes de type T ou de type B. (A) : gènes connus comme étant impliqués dans les leucémies B ; (B) : gènes connus comme étant impliqués dans les leucémies T. La couleur **verte** représente les gènes qui sont sous-exprimés dans les leucémies par rapport au contrôle et la couleur **rouge** indique ceux qui sont sur-exprimés dans les leucémies par rapport au contrôle. Les valeurs sont exprimées en logarithme base 2. Ainsi une valeur de 4 signifie une variation d'expression de 2^4 et donc de 16 fois.

3.2.2.4 Etude de la variation intra-leucémique de l'expression de certains gènes

Dans le but de déterminer l'éventuelle présence de variations au niveau de l'expression de certains gènes entre les sous-groupes de leucémies (entre T1, T2 et T3 et entre B1, B2 et le B3), nous avons utilisé deux approches informatiques : RMA et une autre qui est plus visuel et qui est nommé MAS5.

3.2.2.4.1 Analyse par RMA

Afin d'étudier ces variations intra-leucémiques et grâce au programme RMA, nous avons comparé le niveau d'expression de chaque leucémie par rapport à son contrôle correspondant. Le «fold-change» que nous avons utilisé était de 2. Comme le montre le tableau 3.5, nous avons trouvé des gènes qui étaient spécifiques à chaque sous- type de leucémie : certains étaient spécifiques à T1 à T2 ou à T3, alors que d'autres étaient spécifiques à B1 à B2 ou à B3 (pour plus de détails, voir le tableau 3.5). En comparant les valeurs qui se trouvent au niveau de la dernière colonne pour chaque type de leucémie, nous avons remarqué que la leucémie T3 présente le plus grand nombre de gènes différenciellement exprimés par rapport au contrôle; T par contre, la leucémie T1 en présente le moins. Ceci est logique étant donné qu'il s'agit du même linéage pour le contrôle T et la leucémie T1. De son côté, la leucémie B2 exprime plus de gènes qui varient par rapport au contrôle B alors que la leucémie B3 en exprime le moins.

Tableau 3.5 Nombre de gènes exprimés en fonction du type de leucémie

	Gènes spécifiques d'un sous-type de leucémie		
	Sur-exprimés par rapport au contrôle	Sous-exprimés par rapport au contrôle	Total des gènes spécifiques à chaque sous-type
T1 (CD4 ⁺ /CD8 ⁺)	219	119	338
T2 (CD4 ⁺ /CD8 ⁺)	513	249	762
T3 (CD4 ⁺ /CD8 ⁺)	972	553	1525
B1 (CD45R ^{low} /CD19 ⁺)	385	247	632
B2 (CD45R ⁺ /CD19 ⁺)	508	494	1002
B3 (CD45R ⁺ /CD19 ⁺ /SCA1 ⁺)	305	268	573

Afin de sélectionner ces gènes, nous avons appliqué le programme RMA. Les paramètres choisis étaient que ces gènes devaient varier d'un minimum de 2 fois entre le-sous type de leucémie en question. Par contre, ces gènes ne devaient pas varier plus de 1.5 fois pour le reste des leucémies. Nous avons ainsi réussi à trouver des variations intra-leucémiques. La dernière colonne représente la somme de ces deux catégories. Les chiffres en gras représentent les leucémies possédant le plus de gènes différentiellement exprimés par rapport aux contrôles.

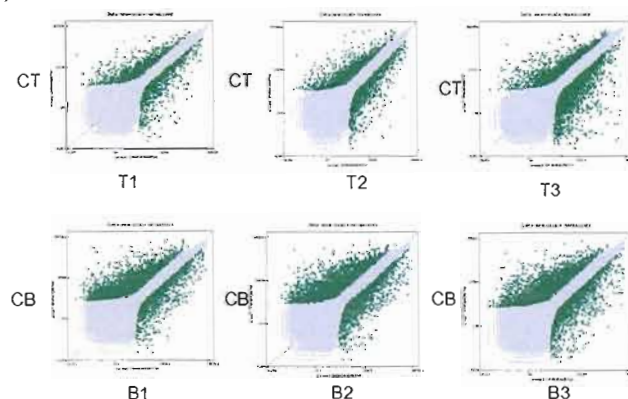
3.2.2.4.2 Analyse par «Scatter plots MAS 5.0» (diagrammes de dispersion)

Le diagramme de dispersion est une autre méthode simple que nous avons appliquée et qui nous a permis de mieux visualiser les variations intra-leucémiques. Ce programme est disponible gratuitement sur le site Affymetrix. Grâce à ce programme, il nous était possible de comparer les leucémies deux à deux. Chaque point vert représente un gène qui varie entre les deux échantillons à comparer. Comme le montre la figure 3.9. A, le nombre le plus élevé de gènes différentiellement exprimés entre les leucémies T et le CT est obtenu pour la leucémie T3 (voir la masse de points verts). D'un autre côté, nous pouvons remarquer que la leucémie B2 présente le plus de gènes différentiellement exprimés par rapport au CB lorsque nous la comparons à B1 et à B2.

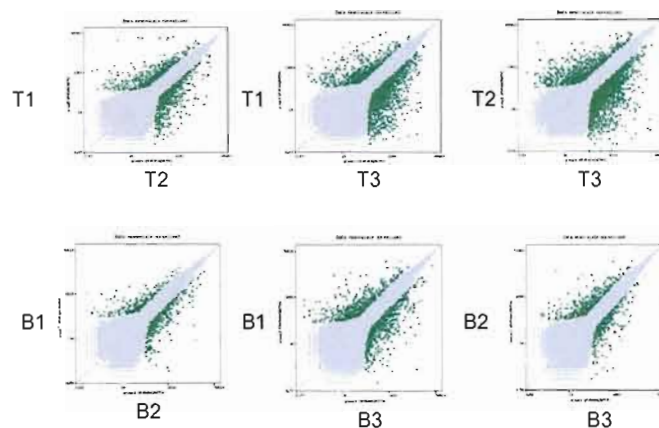
Etant donné que ce test est simple à réaliser, nous avons aussi voulu comparer les leucémies deux à deux. Donc, comme le montre la figure 3.9. B, nous pouvons d'abord remarquer qu'il existe beaucoup moins de gènes différentiellement exprimés entre les leucémies du même type (figure 3.9.B) qu'entre les leucémies et leur contrôle (figure 3.9.A). Dans cette figure, nous pouvons voir que la différence d'expression est plus importante en comparant la leucémie T2 avec la leucémie T3 et en comparant la leucémie B1 avec la leucémie B3.

Dans la figure 3.9. C, nous avons comparé simultanément les 3 leucémies T au CT et les 3 leucémies B au CB. Cette comparaison est identique à celle que nous avons réalisée à l'aide du programme RMA (Voir tableau 3.3) pour déterminer les gènes spécifiques aux leucémies T et ceux qui sont spécifiques aux leucémies B. Ainsi, en observant le diagramme de dispersion de la figure 3.9. C, nous pouvons voir que, par rapport aux leucémies T, les leucémies B présentent un plus grand nombre de gènes différentiellement exprimés par rapport au CB.

(A)



(B)



(C)

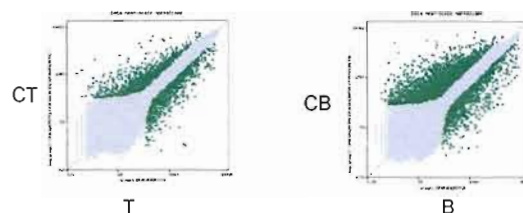


Figure 3.9 Comparaisons intra- et inter-leucémique à l'aide du diagramme de disparition. Grâce au programme MAS5, nous avons réalisé différents types de comparaisons que ce soit entre leucémies d'un même type ou entre une leucémie et son contrôle. Chaque point vert représente un gène. Les points qui sont collés ou proches de la médiane (la partie colorée en gris), représentent les gènes qui ne varient pas entre les deux échantillons à comparer. Par contre, ceux qui sont plus éloignés de cette médiane et qui se trouvent du côté de l'axe des Y représentent les gènes qui sont présents au niveau de l'échantillon situé au niveau de l'axe des Y. Il en va de même pour les points qui se trouvent éloignés de la médiane et positionnés du côté de l'axe des X. (A) : Comparaison de chacun des sous-types de leucémies avec le contrôle correspondant ; (B) : Comparaison des sous-types de leucémies d'un même type entre eux ; (C) : Comparaison des 3 leucémies T avec le contrôle T et des 3 leucémies B avec le contrôle B.

3.3 Validation des résultats obtenus suite à l'analyse par les micropuces par RT-PCR

3.3.1 Choix des gènes à analyser

Suite à la recherche bibliographique, afin de valider les résultats des micropuces et parmi l'ensemble des 56 gènes, nous avons sélectionné les gènes qui nous semblaient les plus intéressants (étant donné que ce projet est à long terme). Ceci nous a permis de retenir 7 gènes. Parmi ces derniers, quelques-uns étaient communs aux deux types de leucémies T et B alors que certains étaient spécifiques à chaque type de leucémie (tableau 3.6). Dans le tableau 3.6, toutes les valeurs exprimées sont sous la forme d'un log base 2.

Tableau 3.6 Niveau d'expression des 7 gènes sélectionnés dans les différentes leucémies d'après les micropuces à ADN

(A)

ProbesetIDs	T1-CT	T2-CT	T3-CT	B1-CB	B2-CB	B3-CB	Gene Symbol
1419122_at	2.333	2.162	2.846	-0.4098	-0.069	-0.152	Mettl1
	5.03	4.47	7.19	-1.32	-1.04	-1.11	
1424981_at	2.947	3.602	2.596	-0.3606	-0.406	-0.297	Nln
	7.71	12.14	6.04	-1.28	-1.32	-1.22	
1429688_at	-0.48	-0.125	-0.377	3.2368	2.7795	2.7528	Arntl2
	-1.39	-1.09	-1.29	9.42	6.86	6.74	
1456741_s_at	0.065	-0.036	-0.06	4.0943	3.9247	3.2507	Gpm6a
	1.04	-1.02	-1.04	17.07	15.17	9.51	
1440633_at	2.874	3.416	1.682	3.6951	3.9754	3.5176	Neil3
	7.33	10.62	3.2	12.95	15.72	11.44	
1418925_at	2.376	1.505	1.196	2.9873	2.9929	2.409	Celsr1
	5.19	2.82	2.29	7.88	7.94	5.27	
1454654_at	-3.25	-4.03	-4.488	-1.1304	-1.716	-1.471	Dir2
	-9.51	-16.33	-22.43	-2.17	-3.28	-2.77	

(B)

CT	T1	T2	T3	CB	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3	CE	Gene Symbol
-1.569	0.764	0.593	1.271	0.0101	-0.4	-0.059	-0.142	0.1027	-0.114	0.376	0.048	0.1081	-0.014	-0.077	-0.904	Mettl1
-2.96	1.69	1.5	2.42	1	-1.31	-1.04	-1.1	1.07	1.08	1.29	1.39	1.07	-1	-1.7	-1.87	
-1.845	1.102	1.757	0.751	-0.139	-0.5	-0.545	-0.436	0.5374	-0.187	0.403	0.201	-0.386	0.0267	-0.461	-0.939	Nln
-3.59	2.14	3.37	1.68	-1.1	-1.4	-1.45	-1.35	1.45	-1.13	1.32	1.81	-1.3	1.2	-1.37	-1.91	
-0.367	-0.849	-0.49	-0.744	-0.382	2.855	2.3978	2.3709	-0.616	-0.529	-0.636	-0.529	-0.524	-0.827	-0.695	-0.634	Arntl2
-1.28	-1.8	-1.4	-1.67	-1.3	7.23	5.26	5.17	-1.53	-1.44	-1.55	-1.44	-1.43	-1.54	-1.61	-1.55	
-0.781	-0.716	-0.82	-0.841	-0.425	3.6897	3.5001	2.8281	-0.957	-0.907	-0.837	-0.67	-0.989	-0.793	-0.584	-0.68	Gpm6a
-1.71	-1.64	-1.76	-1.79	-1.34	12.71	11.31	7.09	-1.94	-1.87	-1.78	-1.59	-1.98	-1.73	-1.94	-1.6	
-1.996	0.878	1.42	-0.314	-2.493	1.202	1.4823	1.0245	-0.408	0.879	-0.437	-0.017	-0.82	-1.281	-0.934	1.814	Neil3
-3.98	1.83	2.67	-1.24	-5.62	2.3	2.78	2.03	-1.32	1.83	-1.35	-1.01	-1.76	-2.43	-1.91	3.51	
0.7717	3.148	2.276	1.968	-1.267	1.7198	1.7254	1.1415	-1.013	-1.45	-1.728	-0.9	-1.649	-1.816	-1.619	-1.507	Celsr1
1.7	8.86	4.82	3.91	-2.4	-3.29	3.3	2.2	-2.01	-2.73	-3.31	-1.89	-3.13	-3.06	-3.07	-2.84	
3.0646	-0.19	-0.96	-1.423	0.0503	-1.08	-1.665	-1.421	3.8281	-1.555	0.734	-0.518	-1.185	0.964	0.0319	1.499	Dirc2
8.33	-1.14	-1.94	-2.68	1.03	-2.11	-3.17	-2.67	12.36	-2.93	1.66	-1.43	-2.27	1.99	1.02	2.82	

La couleur **rouge** correspond aux gènes qui sont sur-exprimés dans les leucémies (A) : par rapport aux contrôles ou bien (B) : par rapport à l'ensemble de la moyenne des leucémies alors que la couleur **verte** indique ceux qui sont sous-exprimés. Les valeurs non colorées représentent les gènes qui ne varient pas significativement au niveau des leucémies (A) : par rapport à leur contrôle ou bien (B) : par rapport à l'ensemble de la moyenne des leucémies (la variation ne doit pas dépasser les 1.5 fois). Les valeurs indiquées dans le tableau sont sous la forme logarithmique et pour retrouver la vraie valeur de l'expression différentielle d'un gène entre une leucémie et son contrôle, nous devons appliquer la formule suivante : 2^n , où n représente les valeurs inscrites dans le tableau (pour faciliter la compréhension de ce tableau, la conversion des valeurs logarithmiques a été réalisée au niveau des lignes colorées en **bleu**). Par exemple pour le gène *Mettl1*, la valeur logarithmique de T1-CT correspond à 2.333. La conversion de ce chiffre est situé au niveau de la ligne bleu, juste en dessous de ce nombre et correspondant à 5.03).

Comme le montre ce tableau :

- les gènes *Mettl1* et *NLN* qui sont spécifiques aux trois leucémies T, sont sur-exprimés d'un minimum de 4 fois par rapport au contrôle T. La valeur la plus basse est celle présente au niveau de T2 pour le gène *Mettl1* où il est sur-exprimé de 4.47 fois par rapport au CT. Nous remarquons que, pour ces gènes, la variation entre les 3 leucémies B par rapport au CB est très faible. Elle se situe entre 1.04 (gène *Mettl1*, leucémie B2) et 1.3 (gène *Mettl1*, la leucémie B1) (tableau 3.6 A).
- Les gènes *Arntl2* et *Gpm6a* qui sont spécifiques aux leucémies B, sont sur-exprimés d'un minimum de 6 fois par rapport au contrôle B. La valeur la plus basse est celle présente au niveau de B3 pour le gène *Arntl2* où il est sur-exprimé de 6.74 fois par rapport au CB. De même, ce tableau montre que les 3 leucémies T ne varient presque

pas par rapport au CT. Cette faible variation se situe entre 1.04 (gène *Gpm6a*, leucémie T3) et 1.3 (gène *Arntl2*, leucémie T3) (tableau 3.6 A).

- Les gènes *Neil3* et *Celsr1* qui sont communs aux leucémies T et B, sont sur-exprimés par rapport à leur contrôle respectif. Cette sur-expression varie entre 2.3 (gène *Celsr1*, leucémie T3) et 15.2 fois (gène *Neil3*, leucémie B2) (tableau 3.6 A).
- Finalement le gène *Dirc2* qui est commun aux deux types de leucémies est sous-exprimé, par rapport à leur contrôle correspondant, d'un minimum de 2.2 (leucémie B2) et d'un maximum de 22.4 fois (leucémie T3) (tableau 3.6 A).

Le tableau 3.6.B montre le résultat obtenu suite à la comparaison de l'expression des gènes dans chaque échantillon avec la moyenne d'expression de chaque gène au niveau de tous les échantillons. Ce tableau nous a permis de confirmer les résultats du tableau 3.6.A malgré le fait que nous ayons utilisé deux raisonnements différents. Pour les gènes *Mettl1* et *Nln*, nous pouvons voir qu'ils sont fortement exprimés au niveau des leucémies T. Par contre et pour le reste des leucémies, leur expression ne varie pas significativement par rapport à la moyenne (sauf pour le gène *Nln* dans la leucémie E3). Pour les gènes *Arntl2* et *Gpm6a*, nous remarquons qu'ils sont très fortement exprimés au niveau des leucémies B. Par contre, au niveau du reste des échantillons, leur expression est faible ou peu variable par rapport à la moyenne. Ce tableau nous montre que contrairement au reste des gènes situés ci-dessus, *Neil3* a une expression plus hétérogène un peu différente de celle présentée dans le tableau 3.6.A. Ce gène est fortement exprimé au niveau des leucémies T à l'exception de T3 et des leucémies B. Il l'est aussi au niveau de la leucémie E1 et au niveau du contrôle E. Par contre, pour le reste des échantillons, nous avons remarqué que le niveau d'expression de *Neil3* a soit diminué soit resté constant par rapport à la moyenne. Pour le gène *Celsr1*, nous pouvons voir que par rapport à la moyenne, il se trouve fortement exprimé au niveau des leucémies T et B (malgré qu'il l'est aussi pour le CT mais en moindre proportion) et faiblement exprimé au niveau du reste des échantillons. Finalement, et comme pour le gène *Neil3*, *Dirc2* a un niveau d'expression assez hétérogène et est faiblement exprimé au niveau des leucémies lymphoïdes (excluant T1) ainsi qu'au niveau des leucémies E1 et Mk1. Par contre, nous avons remarqué

qu'au niveau du reste des échantillons et par rapport à la moyenne, son niveau d'expression est resté faible pour certains et peu variable dans d'autres.

3.3.2 Validation du résultat des micropuces à ADN par RT-PCR

Pour valider les résultats obtenus suite à l'analyse par la technique des micropuces à ADN et pour l'optimisation de la réaction de RT-PCR (non semi-quantitative), nous avons d'abord testé l'expression de certains des gènes sélectionnés au niveau de leucémies non triées, lymphoïdes et non lymphoïdes. Ces leucémies sont spécifiées dans le tableau 3.7 :

Tableau 3.7 Leucémies non triées

T4	T5	T6	Souris contrôle C	B4	B5	B6	E4	Mk4	M4
57.4	61.7	57.1		43.2	6.1	67.2	29.4	10.6	04.2
CD4 ⁺ CD8 ⁺ et CD4 ⁻ CD8 ⁺	CD8 ⁺ CD4 ⁻	CD4 ⁺ CD8 ⁻	Rate d'une souris normale	CD45R ⁺ CD71 ⁺ CD19 ⁺			Ter-119 ⁺ CD71 ⁺	CD41 ⁺ c-kit ⁺	Mac1 ⁺ Gr1 ⁻

Afin d'étudier le niveau d'expression de certains gènes nous avons sélectionné 10 échantillons que nous avons préalablement immunophénotypés par cytométrie de flux. Ces échantillons sont constitués de 3 leucémies de type T (T4, T5 et T6), de 3 leucémies de type B (B4, B5 et B6), d'un échantillon contrôle provenant de la rate d'une souris normale et finalement de 3 leucémies non lymphoïdes (E4, Mk4 et M4). Les phénotypes de ces leucémies sont indiqués au niveau de la dernière rangée du tableau.

La figure 3.10 montre que, pour chacun des gènes étudiés ainsi que pour chaque paire d'amorces, la leucémie T4 semble dégradée. Il serait donc plus prudent de ne pas la prendre en considération lors de cette analyse.

Dans la figure 3.10. A, nous avons étudié le niveau d'expression des gènes spécifiques aux leucémies T (*Mettl1* et *NLN*). Comme le montre le graphique de cette figure, nous remarquons que :

- l'expression du gène *Metll1* est relativement faible sur l'ensemble des échantillons testés. Ce niveau d'expression reste à peu près semblable pour le reste des échantillons.
- pour le gène *NLN*, les leucémies T5 et T6 présentent le plus haut niveau d'expression par rapport aux restes des leucémies et même par rapport à l'échantillon contrôle.

Dans la figure 3.10. B, nous avons testé le niveau d'expression du gène *Gpm6a* (paire 1 ou séquence Affymétrie) qui est spécifique aux leucémies B. Nous avons trouvé que la leucémie B4 et un peu moins la leucémie B6, présentent le plus haut niveau d'expression. En observant le gel d'agarose, nous pouvons voir que l'échantillon pour la leucémie B5 est partiellement dégradé ce qui explique son bas niveau d'expression. EN plus nous pouvons constater que pour cette leucémie, le poids moléculaire de cette bande est plus haut que celui de tous les autres échantillons. Ceci pourrait être expliqué par une contamination à l'ADN. Cette figure nous montre aussi que le niveau d'expression de ce gène au niveau des leucémies M2 et Mk4, est quand même assez élevé. Cependant, cette expression reste faible pour le reste des échantillons (T5, T6, contrôle et E4).

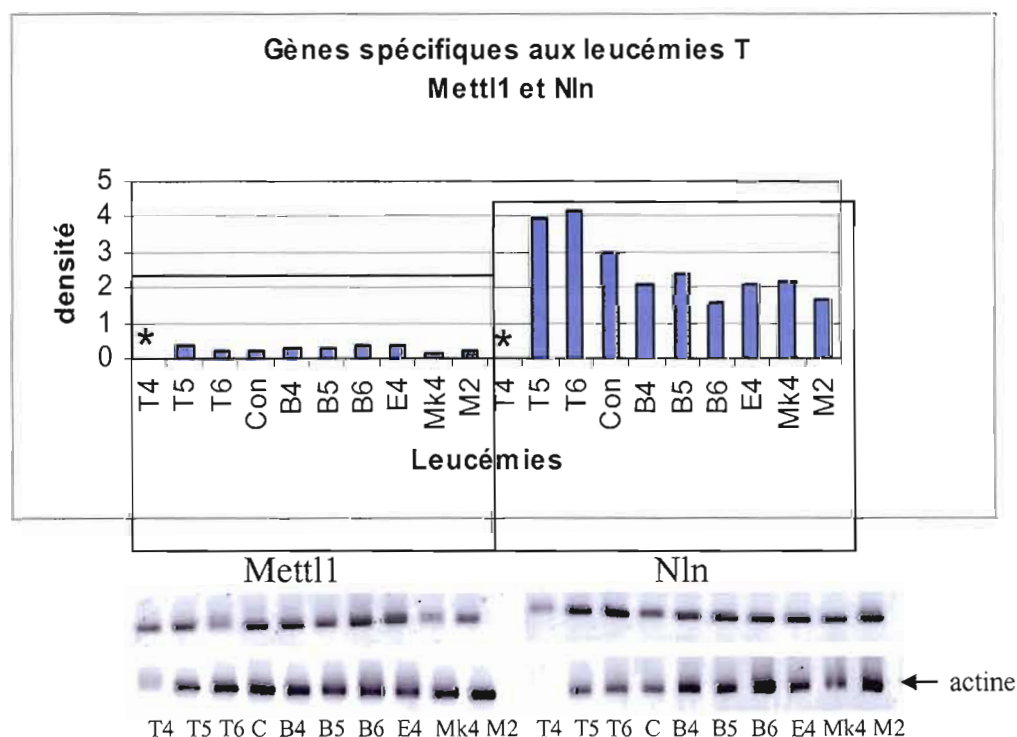
Dans la figure 3.10. C se trouvent les gènes qui sont communs aux leucémies T et B et sur-exprimés par rapport au contrôle T et B correspondant. En observant le graphique, nous pouvons remarquer que :

- Pour la première paire du gène *Neil3* (*Neil3.1*) englobant les deux exons qui encadrent la séquence Affymetrix et correspondant donc au seul transcrit de ce gène (disponible sur le site Ensembl), le niveau d'expression de ce gène est plus élevé au niveau des leucémies T5, T6 et B5 par rapport au reste des échantillons. Cependant, pour l'ensemble des leucémies lymphoïdes, seule T5 se trouve sur- exprimée par rapport au contrôle.
- Pour la deuxième paire du gène *Neil3* (*Neil3.2*) représentant la séquence Affymétrie et correspondant à l'EST qui serait un autre transcrit de ce gène (présente dans le site UCSC), la variation d'expression entre les différents échantillons n'est pas très significative excluant les leucémies T6 et B5 pour lesquelles nous pouvons noter une légère sur-expression par rapport à l'ensemble des échantillons.

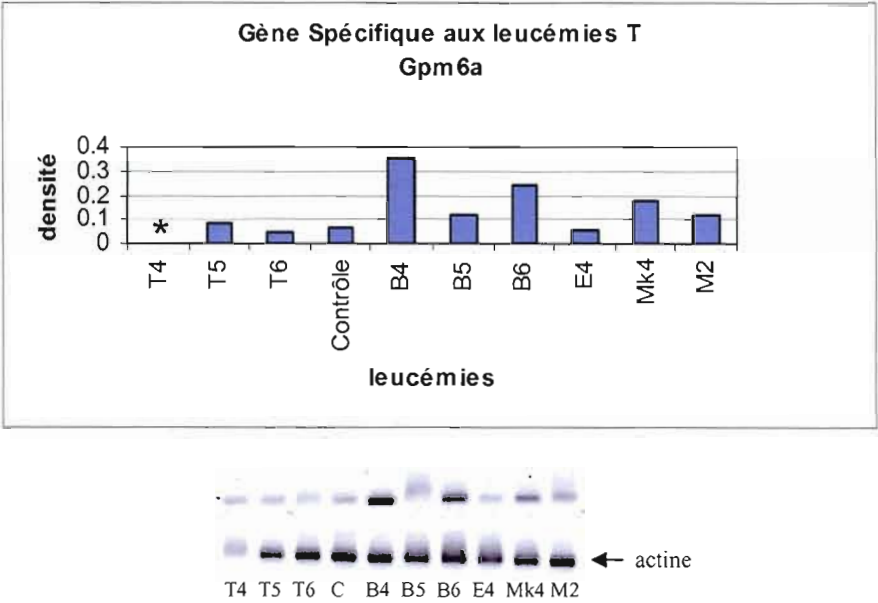
- Malgré qu'il parait très difficile de voir les différences entre les leucémies lymphoïdes et celles non lymphoïdes avec l'échelle de densité de la figure 3.10. C, nous pouvons quand même voir que le niveau d'expression du gène *Celsr1* est légèrement plus élevé au niveau des leucémies T5, T6 et B4 par rapport aux leucémies non lymphoïdes. Cette différence n'est pas plus prononcée en comparant l'expression de *Celsr1* entre ces leucémies lymphoïdes et le contrôle. Toutefois, nous pouvons noter que B5 est la seule leucémie lymphoïde qui ne varie pas significativement par rapport au contrôle.

Pour les cellules non triées, nous avons aussi étudié le niveau d'expression des gènes *Arntl2* et *Dirc2* (*Dirc2.1* et *Dirc2.2*). Cependant, les bandes que nous avons obtenues que ce soit pour ces gènes ou pour l'actine correspondante étaient diffuses. C'est pour cette raison que nous n'avons pas pu les intégrer dans ce travail.

(A)



(B)



(C)

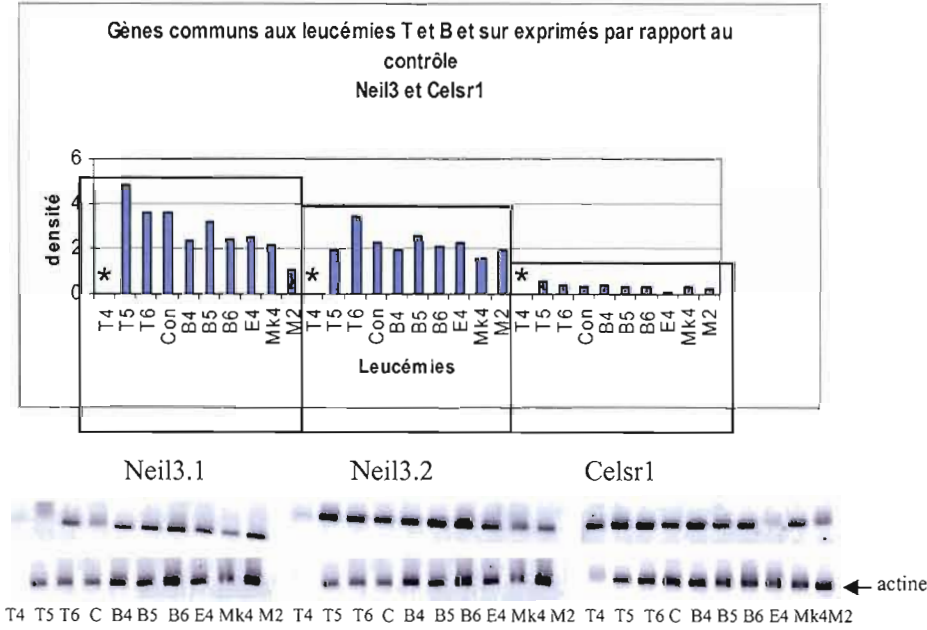


Figure 3.10 Etude de l'expression de quelques gènes, par RT-PCR, au niveau des leucémies non triées. Pour chacun des gènes présentés dans cette figure, nous avons déterminé le profil d'expression au niveau de 10 échantillons (3 leucémies T, 3 leucémies B, un échantillon provenant d'une rate normale, 1 leucémie E, 1 leucémie Mk et une leucémie M). En parallèle, afin de normaliser nos résultats, nous avons déterminé le niveau d'expression de l'actine. Ceci nous a permis de construire l'ensemble des 3 graphiques situés au-dessus de chacun des gels. (A) : Gènes spécifiques aux leucémies lymphoïdes de type T (*Mettl1* et *NLN*); (B) : Gènes spécifiques aux leucémies lymphoïdes de type B (*Gmp6a.1*); (C) : Gènes communs aux leucémies T et B et sur exprimés par rapport aux contrôles B et T correspondants (*Neil3.1*, *Neil3.2* et *Celsr1*). Le symbole étoile (*) signifie que le fragment est dégradé.

Remarques

- pour la figure 3.10, il nous est impossible de représenter les barres d'erreurs étant donné que cette expérience n'a été réalisée qu'une seule fois. Il serait donc nécessaire de refaire l'expérience deux ou trois fois pour mieux valider nos résultats.
- d'après cette figure, nous pouvons remarquer que la différence d'expression d'un gène entre l'ensemble des échantillons analysés n'est pas aussi prononcée que celle obtenue suite à l'analyse par la technique des micropuces à ADN. Ceci est dû au fait que la validation des résultats des micropuces n'a pas été réalisée à l'aide de la RT-PCR quantitative. Cette dernière devrait ainsi permettre de bien distinguer les plus petites variations entre les différents échantillons analysés.

Suite à ce premier test d'expression des 7 gènes sur des leucémies non triées, nous les avons testés au niveau de leucémies triées. Ces dernières sont présentées dans le tableau 3.8 :

Tableau 3.8 Ensemble des leucémies triées

B7	B8	CB2	T7	T8	CT2
67.2 (NFS)	67.4 (NFS)	B-C01 (FVB)	61.3 (NFS)	61.7 (NFS)	T-Co1 (FVB)
CD45R ⁺ CD19 ⁺ CD71 ⁺	CD45R ⁺ CD71 ⁺ CD19 ⁺	CD45R ⁺ CD71 ⁺ CD19 ⁺	CD4 ⁺ CD8 ⁺	CD8 ⁺ CD4 ⁺	CD4 ⁺ CD8 ⁺

Afin d'étudier le niveau d'expression des 7 gènes que nous avons sélectionnés pour valider les résultats des micropuces à ADN, nous avons choisi 6 échantillons constitués de 2 leucémies de type B (B7, B8), de 2 leucémies de type T (T7, T8) et d'un échantillon contrôle provenant de la rate d'une souris normale pour chaque type de leucémie (CB2 et CT2). Les phénotypes de ces leucémies sont indiqués au niveau de la dernière ligne de ce tableau.

Dans la figure 3.11, nous pouvons remarquer que le contrôle B est partiellement dégradé pour l'ensemble des 7 gènes testés. Malheureusement, il nous était impossible d'obtenir un autre échantillon contrôle B étant donné que nous ne disposions plus du nombre suffisant de souris (12 souris normales). C'est d'ailleurs ce même problème qui nous a contraint à utiliser un contrôle B (CB2) provenant de souris FVB au lieu d'un contrôle B provenant de souris NFS (les leucémies proviennent de souris NFS). Par contre, ce dernier problème n'est pas majeur du fait que ces deux souches sont très proches phylogéniquement. Pour l'analyse de nos résultats, il serait plus prudent de ne pas prendre l'échantillon B2 en considération. Nous devrions donc nous fier plus aux résultats obtenus pour les leucémies T et être plus vigilants dans l'interprétation surtout que les signaux de l'actine semblent souvent être très faibles ou dégradés.

Dans la figure 3.11. A, se trouvent les gènes qui sont spécifiques aux leucémies T :

- En comparant le niveau d'expression du gène *Mettl1* entre les leucémies T et le CT2, nous remarquons que ce gène est environ 3 fois sur-exprimés dans les leucémies T7 et T8. Nous avons aussi trouvé que *Mettl1* est fortement exprimé au niveau des leucémies B. Étant donné que l'échantillon CB2 est dégradé, nous pouvons supposer qu'il se trouve exprimé, à peu près, au même niveau que les leucémies B (si nous nous basons sur les résultats du tableau 3.5).

- Le gène *Nln* semble aussi fortement exprimé au niveau des leucémies T7 et T8 et non exprimé au niveau du CT2. De même que *Mettl1*, et malgré que *Nln* se trouve tout aussi fortement exprimé au niveau des leucémies B, nous ne pouvons pas conclure quant au niveau d'expression de ce gène au niveau du CB2. Cependant, comme indiqué le tableau 3.5, nous pouvons dans ce cas aussi, supposer qu'il n'y pas de variation au niveau de l'expression de *Nln* entre le CB2 et les leucémies B.

Dans la figure 3.11. B, sont représentés les gènes qui sont spécifiques aux leucémies B :

- Pour la première paire d'amorce du gène *Arntl2* englobant une partie de la séquence Affymétrie et du dernier exon du transcrit de ce gène ainsi que pour la deuxième paire d'amorce du gène *Arntl2* englobant la séquence Affymétrie, nous pouvons observer que le gène *Arntl2.1* semble uniquement exprimé au niveau des leucémies B. Cependant nous ne pouvons conclure pour le CB2. Toutefois d'après le tableau 3.5, nous pouvons supposer que pour les leucémies B, ce gène se trouve sur-exprimé par rapport au CB2. Pour la séquence Affymétrie correspondant à ce gène, il est intéressant de noter, comme mentionner précédemment, qu'elle est située au niveau de la partie 3'UTR du transcrit. Cependant dans la banque donnée NCBI, cette région n'a pas encore été séquencée. Ainsi le fait de positionner un oligo sens au niveau du dernier exon du transcrit du gène *Arntl2* et un oligo anti-sens au niveau de la séquence Affymétrie, nous a permis d'obtenir une bande d'environ 800 pb. Qui représenterait une partie du 3'UTR de *Arntl2*.
- Pour la première paire d'amorce du gène *Gpm6a* (*Gpm6a.1*) située au niveau de la séquence Affymétrie, nous avons observé le même résultat que celui du gène *Arntl2*. En testant la deuxième paire d'amorce de gène *Gpm6a* (*Gpm6a.2*) qui est située au niveau des deux exons du transcrit de ce gène (résultat non montré), nous n'avons eu aucune amplification, ni pour les leucémies ni pour les contrôles (nous avons obtenu le même résultat que pour les leucémies non triées).

La figure 3.11. C inclut les gènes qui sont communs aux leucémies T et B et sur-exprimés par rapport aux contrôles T et B correspondants :

- La première paire d'amorces du gène *Neil3*, notée *Neil3.1*, est positionnée au niveau de deux exons du seul transcrit de ce gène offert par le site Ensembl. Cette figure nous permet de voir que les leucémies B7, B8 et T8 ont un haut niveau d'expression. Par contre, nous remarquons que ce gène semble se trouver moins exprimé au niveau de la leucémie T7. Ce résultat corrèle bien avec le tableau 3.6.B où la leucémie T3, identique à la leucémie T7 de la figure 3.11.C (toutes les deux sont CD4⁺/CD8⁺), ne varie pas d'une façon significative par rapport à la moyenne en comparaison avec les deux autres leucémies de ce tableau. Pour les leucémies T, seule T8 est sur-exprimée par rapport au CT2.
- La deuxième paire d'amorces du gène *Neil3* est notée *Neil3.2*. Cette dernière est positionnée au niveau de la séquence Affymetrix qui, selon le programme UCSC, serait contenue au niveau d'un exon d'une EST correspondant à *Neil3* et nommée AK031027. Pour cette paire, nous avons obtenu le même résultat que pour la première paire. Cependant, le niveau d'expression du gène *Neil3* paraît devenir presque identique à celui du CT2.
- En observant les résultats du gène *Celsr1*, nous pouvons voir une légère expression au niveau de la leucémie B8. Par contre, aucune amplification n'a été observée pour la leucémie B7. Cependant en comparant les leucémies T avec le CT2, il semble que ce gène soit 2 fois plus exprimé au niveau de la leucémie T7 et 3 fois plus exprimé au niveau de la leucémie T8 par rapport au CT2.

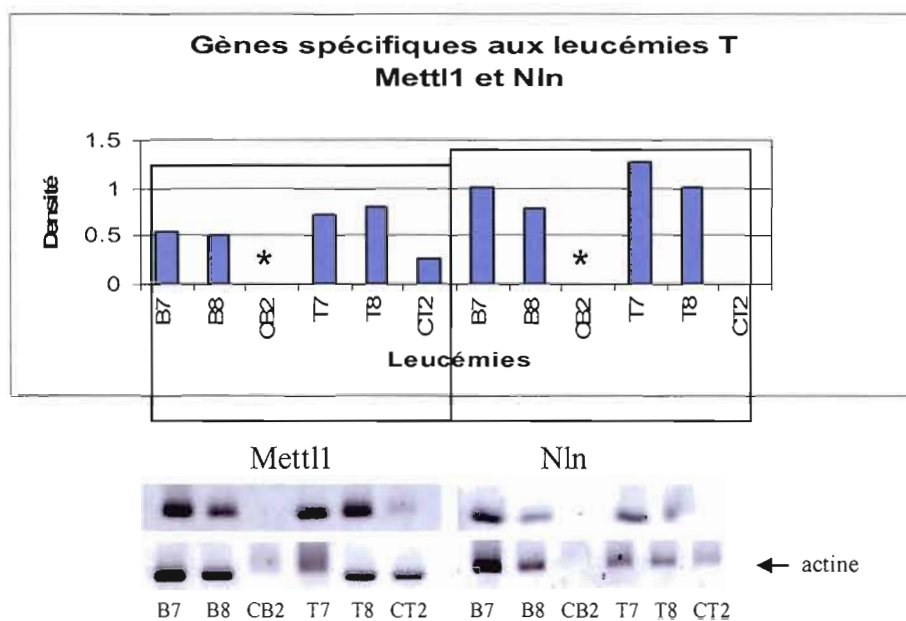
Finalement, la figure 3.11. D représente le niveau d'expression du gène *Dirc2* qui est commun aux leucémies T et B et est sous-exprimé par rapport aux contrôles T et B correspondants :

- La première paire d'amorces de ce gène permet l'amplification du transcrit du gène *Dirc2* (*Dirc2.1*). D'après le graphique, nous pouvons constater que le niveau d'expression du gène *Dirc2.1* est un peu élevé au niveau de la leucémie B7 et B8. Etant donnée que le CB2 est dégradé, nous pouvons supposer qu'en réalité, *Dirc2* est beaucoup plus exprimé au niveau de ce contrôle et que cette expression dépasse celle des leucémies B. Ce même graphique montre que, par rapport au CT2,

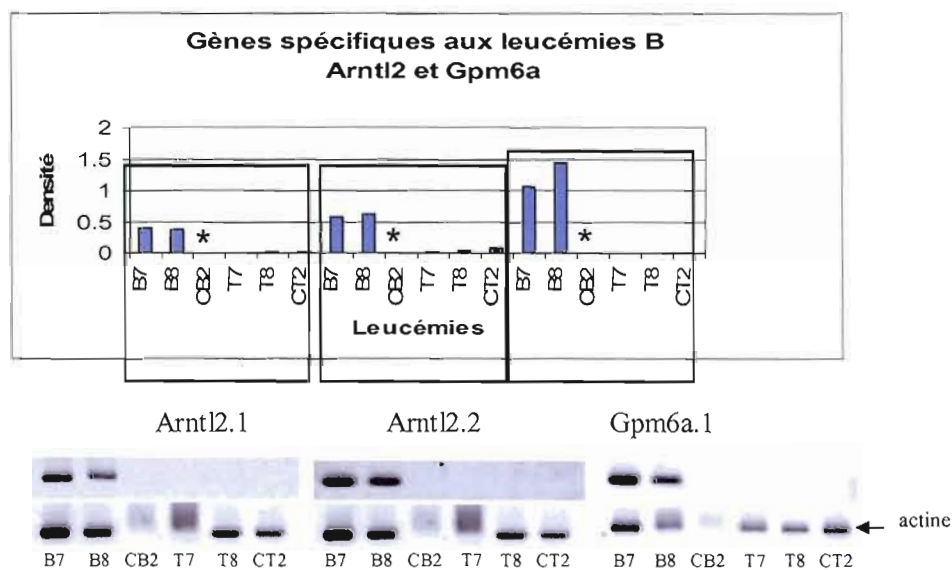
ce gène est 4 fois moins exprimé dans la leucémie T7. Cependant, nous pouvons voir que cet échantillon est aussi partiellement dégradé. Par contre, au niveau de la leucémie T8, il paraît un peu plus exprimé en comparaison avec le CT2.

- La deuxième paire d'amorces de ce gène permet l'amplification de la séquence Affymetrix du gène *Dirc2* (*Dirc2.2*). Nous pouvons noter que le niveau d'expression de ce gène est plus ou moins élevé au niveau des leucémies B7 et B8. Toutefois nous ne pouvons pas conclure s'il se trouve sous ou sur exprimé dans les leucémies B par rapport au CB2. Ce même gène se trouve moins exprimé au niveau des leucémies T7 et T8 par rapport au CT2.

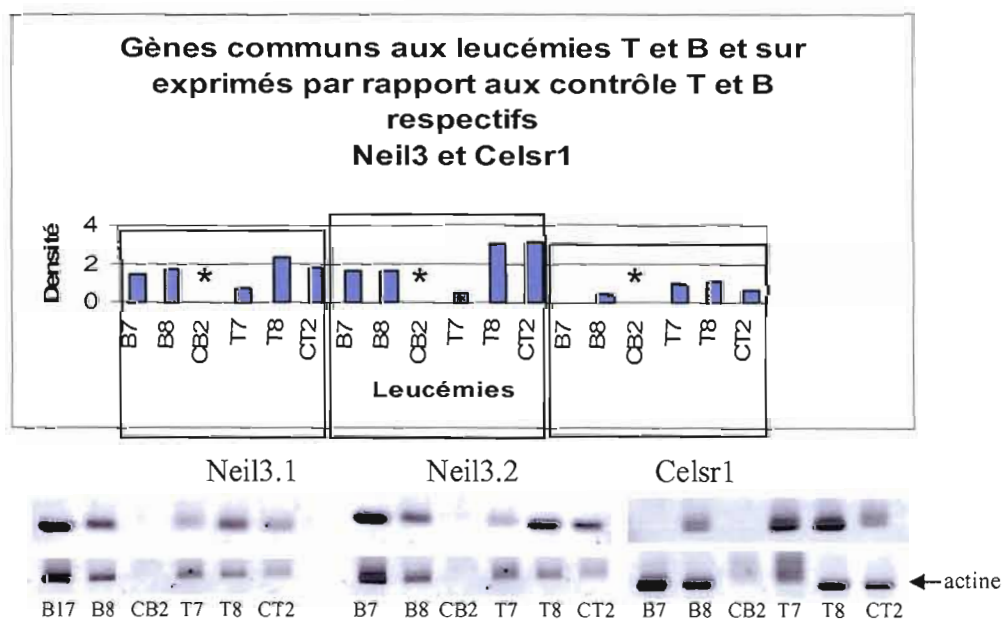
(A)



(B)



(C)



(D)

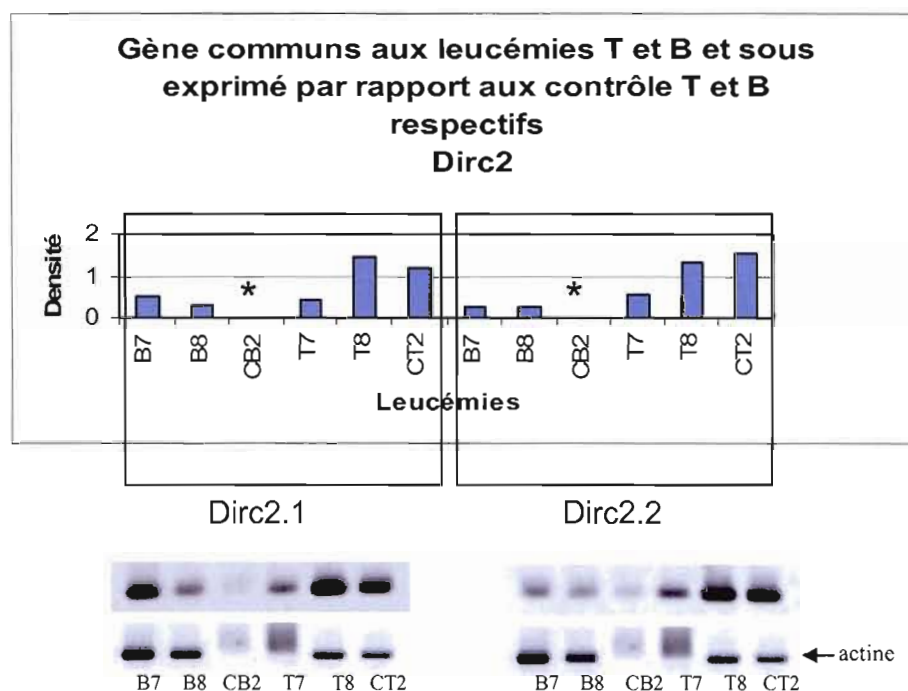


Figure 3.11 Etude de l'expression des 7 gènes sélectionnés au niveau de leucémies triées. Nous avons étudié le niveau d'expression de ces gènes en plus de celui de l'actine, par RT-PCR au niveau de 6 échantillons (2 leucémies B, 2 leucémies T, un contrôle pour les B et un autre pour les T). La normalisation avec le gène référence nous a permis de construire l'ensemble des graphiques situés au dessus des gels. (A) : Gènes spécifiques aux leucémies lymphoïdes de type T (*Mettl1* et *NLN*); (B) : Gènes spécifiques aux leucémies lymphoïdes de type B (*Arntl1.1*, *Arntl2.2* et *Gmp6a.1*); (C) : Gènes communs aux leucémies T et B et sur-exprimés par rapport aux contrôles B et T correspondants (*Neil3.1*, *Neil3.2* et *Celsr1*) ; (D) : Gène commun aux leucémies T et B et sous exprimé par rapport aux contrôles B et T correspondants. Le symbole étoile (*) signifie que le fragment est dégradé.

Remarques

- pour la figure 3.11, il nous est impossible de représenter les barres d'erreurs étant donné que cette expérience n'a été réalisée qu'une seule fois. Il serait donc nécessaire de refaire l'expérience deux ou trois fois pour mieux valider nos résultats.
- d'après cette figure, nous pouvons remarquer que la différence d'expression d'un gène entre l'ensemble des échantillons analysés n'est pas aussi prononcée que

celle obtenue suite à l'analyse par la technique des micropuces à ADN. Ceci est dû au fait que la validation des résultats des micropuces n'a pas été réalisée à l'aide de la RT-PCR quantitative. Cette dernière devrait ainsi permettre de bien distinguer les plus petites variations entre les différents échantillons analysés.

CHAPITRE IV

DISCUSSION

Lors de ce projet, nous nous sommes intéressés à la caractérisation des leucémies lymphoïdes de types T et B induites par le rétrovirus murin Graffi. Différentes techniques ont été appliquées. La première consistait à immunophénotyper les leucémies par cytométrie de flux. Celle-ci avait pour but de permettre une meilleure compréhension de la pathogenèse de ce rétrovirus. La seconde était celle des micropuces à ADN qui a servi à déterminer une panoplie de gènes dont l'expression était dérégulée par le rétrovirus murin Graffi. Dans ce chapitre, nous allons discuter toutes les étapes ainsi que les résultats qui ont amené à une meilleure compréhension des leucémies lymphoïdes induites par le rétrovirus murin Graffi ainsi que celles qui nous ont permis, suite à l'analyse de l'expression différentielle de ces leucémies, de passer d'un nombre de transcrits dépassant les 39 000 à un nombre beaucoup plus restreint.

Pour la première partie et comme montré dans le tableau 3.1, les souris Balb/c inoculées avec la même quantité des variants rétroviraux GV-1.2 et GV-1.4, développent la leucémie après une période de latence plus longue pour celles injectées avec GV-1.4. En effet, il est connu que ces deux variants ont une similitude structurale sauf que le clone GV-1.2 nécessite une période de latence plus courte pour induire la pathologie (environ 80 jours contre 150 jours pour le clone GV-1.4) (Ru et al. , 1993 ; Voisin et al. , 2006). Ceci s'explique par l'existence d'une duplication de 60 bp dans la région " enhancer " au niveau de la séquence U3 des LTR du clone GV-1.2.

L'analyse phylogénétique faite sur le rétrovirus Graffi (Voisin et Rassart, 2006) montre qu'il est très homologue au rétrovirus SRS 19-6. Comme ce dernier, le rétrovirus murin Graffi est un virus qui induit un large spectre de leucémies (Granger et al. , 1999). Ces deux rétrovirus ont la capacité de s'attaquer à différentes lignées dérivant de la cellule souche hématopoïétique. Le rétrovirus murin Graffi peut donc induire soit des leucémies lymphoïdes

de type T ou B soit des leucémies non lymphoïdes. Cette variation dépend du variant rétroviral et de la souche de souris utilisées. En effet, en injectant les souris avec GV-1.4, nos résultats montrent que la souche Balb/c développe majoritairement des leucémies lymphoïdes (de type T) alors que la souche FVB/n développe principalement des leucémies non lymphoïdes (majoritairement de type érythroïde). Dans la souche NFS, le virus Graffi induit quant à lui des leucémies lymphoïdes et non lymphoïdes en plus de leucémies mixtes ou bi-phénotypiques.

Pour une meilleure caractérisation des leucémies induites par le rétrovirus murin Graffi, nous avons sélectionné un certain nombre de leucémies et testé par Northern Blot le niveau d'expression du gène *Fog-1*. *Fog-1*, encore appelé Friend of *GATA-1* (Wechsler et al. , 2002), est connu comme étant spécifique et essentiel au développement des lignées E et Mk (Katz et al. , 2002). Ceci corrèle bien avec nos résultats. En effet, la figure 3.2 nous a montré que *Fog-1* est très faiblement exprimé au niveau des leucémies lymphoïdes en comparaison avec les contrôles. Par contre il est fortement exprimé au niveau des leucémies non lymphoïdes et même au niveau des leucémies mixtes. Pour ces dernières, nous avons noté que le niveau d'expression du gène *Fog-1* se trouve plus élevé dans celles ayant une population érythroïde par rapport à celles ayant une population mégakaryocytaire, ce qui laisse penser que *Fog-1* semble un peu plus spécifique aux leucémies E.

GATA-1, *GATA-2* et *GATA-3* sont trois facteurs de transcription faisant partie de la famille GATA. Plusieurs publications ont montré que *GATA-1* s'exprimait au niveau des cellules érythroïdes primitives et définitives, mégakaryocytaires, éosinophiles et au niveau des mastocytes (Ferreira et al. , 2005). Quant à *GATA-2*, il est impliqué dans la maintenance et dans la prolifération des progéniteurs hématopoïétiques ainsi que dans la spécification des cellules mastocytaires et mégakaryocytaires. Pour *GATA-3*, quant à lui, est connu pour son rôle dans le développement de la lignée cellulaire lymphoïde de type T et dans la différenciation des cellules T CD4⁺ en Th2. Ces facteurs sont essentiels pour le déroulement normal de l'hématopoïèse (Ferreira et al. , 2005) et ont aussi été impliqués dans la tumorigénèse. En effet des mutations observées dans le domaine N-terminal de *GATA-1* sont capables d'induire des maladies humaines dont les leucémies mégakaryocytaires aigues

associées au syndrome de Down ou DS-AMKL («Down syndrome related acute megakaryocytic leukemia»). Il a été montré que le patron d'expression de *GATA-1* et de *GATA-2* peut aider à définir des phénotypes distincts des cellules leucémiques suggérant que le niveau d'expression de ces deux facteurs au niveau des cellules leucémiques est en corrélation avec le stade de différenciation des cellules hématopoïétiques (Shimamoto et al. , 1995). De son côté, dans l'un des travaux publiés, *GATA-3* a été clairement détecté au niveau des leucémies T-ALL (leucémie lymphoblastiques aigues des cellules T) indiquant l'importance de son rôle dans la leucémogénèse (Ono et al. , 1997).

Dans le cadre de notre recherche, nous avons trouvé que *GATA-1* est exprimé partout notamment dans les leucémies lymphoïdes de type T malgré que nous avons remarqué que ce gène est le plus faiblement exprimé au niveau de l'ensemble des leucémies EMk et au niveau de la majorité des leucémies Mk. Pour expliquer cette différence avec la littérature, nous avons émis des hypothèses. Etant donné que chaque système possède ses propres spécificités, l'une des hypothèses est basée sur le fait que l'injection du rétrovirus Graffi aux souris, pourrait leur conférer un type de leucémie particulier les distinguant des autres modèles leucémiques. Ainsi, la présence de Graffi serait responsable de la forte expression de *GATA-1* dans pratiquement tous les types de leucémies induites incluant les leucémies lymphoïdes. Une autre hypothèse qui pourrait expliquer la différence trouvée entre la littérature et entre nos résultats pour *GATA-1* et pour *GATA-2* serait en relation avec le fait que l'utilisation de populations non triées puisse fausser le résultat à cause de la présence de cellules normales non lymphoïdes exprimant *GATA-1* au niveau des leucémies lymphoïdes T et B et exprimant *GATA-2* au niveau des leucémies lymphoïdes de type B. Pour une leucémie lymphoïde de type T, il se pourrait que, pour ces cellules normales non lymphoïdes incluant les érythroïdes, l'érythropoïèse s'arrête à un stade de développement assez avancé où *GATA-1* est fortement exprimé et où *GATA-2* est faiblement exprimé (Figure 4.1) (Ferreira et al. , 2005). D'autre part, une leucémie de type B, peut être contaminée par des cellules normales érythroïdes. Pour ses dernières, il se pourrait qu'elles soient à un stade de développement où l'érythropoïèse, s'est arrêtée à un stade très précoce. Ce stade peut être plus précoce que celui des leucémies T où *GATA-1* est moins exprimé au niveau des cellules normales non lymphoïdes alors que *GATA-2* est plus fortement exprimé (Figure 4.1). A ceci s'ajoute le fait

que certains macrophages pénètrent au niveau de la rate en raison de leur caractère fortement infiltrant. Ces macrophages peuvent ainsi contribuer à augmenter l'expression de *GATA-1* dans les leucémies non lymphoïdes.

Lorsque nous nous intéressons uniquement à *GATA-2*, nous remarquons que ce gène est clairement exprimé au niveau des leucémies non lymphoïdes par rapport aux lymphoïdes à l'exception des leucémies B qui expriment quand même ce gène. Pour *GATA-3*, nos résultats montrent que l'expression de ce facteur de transcription est très élevée dans les leucémies T par rapport à tous les autres types de leucémies.

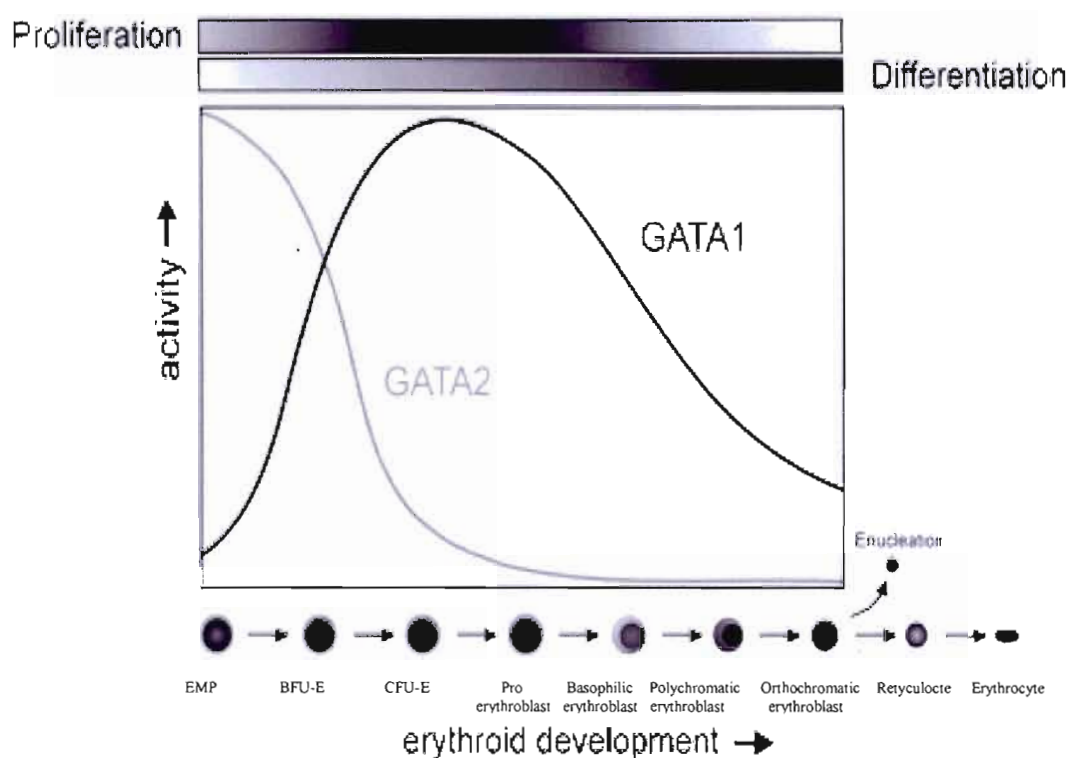


Figure 4.1 Niveau d'expression de GATA-1 et de GATA-2 en fonction de l'érythropoïèse.

Nous avons voulu confirmer les résultats que nous avons obtenus pour *GATA-1* en les comparant au taux d'expression du gène *Fog-1* (spécifique aux leucémies E et Mk), par

analyse de Northern et hybridation avec la sonde spécifique *Fog-1* (travail réalisé par Véronique Voisin) (Annexe B). Nous avons trouvé que l'expression de *Fog-1* est bien faible ou même pas présente au niveau des leucémies lymphoïdes. Par contre, pour les leucémies non lymphoïdes, il existe une grande différence entre le gène *GATA-1* et *Fog-1* allant dans le même sens que l'hypothèse attestant que le rétrovirus murin Graffi induit des leucémies très particulières.

En plus des gènes *GATA-1*, *GATA-2*, *GATA-3* et *Fog-1*, nous avons testé l'expression d'un certain nombre d'autres gènes incluant la myéloperoxydase. Nous avons aussi testé le réarrangement du TCR des lymphocytes T ainsi que celui des chaînes lourdes d'immunoglobulines des lymphocytes B. Malheureusement, nous avons eu de la difficulté lors de l'hybridation des sondes correspondantes.

Pour la deuxième partie de ce projet et afin de déterminer le niveau d'expression des gènes spécifiques ou communs aux leucémies T et B, nous avons trié plusieurs leucémies par la technique des billes magnétiques. La majorité de ces leucémies provient de la souche NFS injectée avec le variant GV-1.4. Ce choix est basé sur le fait que ce modèle induit différents types de leucémies en comparaison avec les deux autres souches de souris (Balb/c et FVB/n) injectées que ce soit avec GV-1.2 ou avec GV-1.4. La seule limitation de notre système était l'apparition d'un pourcentage très élevé de leucémies mixtes (montrant la présence de deux populations leucémiques qui sont différentes l'une de l'autre) et de leucémies bi-phénotypiques (montrant la présence d'une seule population leucémique) (voir annexe A). En dépit de cet inconvénient, le recours à ce modèle a été privilégié du fait que ce projet de recherche est co-mené avec Véronique Voisin dont les préoccupations portent sur l'étude des leucémies non lymphoïdes.

Dans le cadre de ce projet, nous avons déterminé le profil d'expression de 6 leucémies (3 leucémies B et 3 leucémies T) ainsi que celui d'un contrôle pour chacun des deux types T et B. D'autres chercheurs ont déjà étudié le profil d'expression génique pour une meilleure compréhension des leucémies (Jelinek et al. , 2003 ; Ferrando et al. , 2002).

L'avantage de ce projet vient du fait que nous avons la possibilité de comparer les résultats obtenus pour les leucémies lymphoïdes (le présent sujet) avec ceux obtenus pour les leucémies non lymphoïdes (projet de Véronique Voisin). Lors de l'analyse de l'ensemble de nos échantillons par la technique des micropuces à ADN et pour avoir un résultat fiable, l'ARN des échantillons (lymphoïdes et non lymphoïdes) envoyés au Centre Génomique de Montréal a été extrait le même jour dans une période de temps assez courte. Pour les leucémies lymphoïdes, nous avons intentionnellement analysé différents sous-types de leucémies. Par exemple, pour les leucémies T, le sous-type d'un échantillon était $CD4^+/CD8^+$, celui du second était $CD4^+/CD8^-$ et pour le dernier, il était $CD8^+/CD4^-$.

Lors de cette étude, et pour analyser nos résultats, nous avons utilisé différentes méthodes d'analyse. La plus importante était celle nommée Robust Multiple-array Average (RMA). RMA est une méthode spécifiquement créée pour identifier les différences entre les échantillons tout en contrôlant le degré de différence («Fold-change»). Il est connu qu'un large ensemble de données constituées de plusieurs gènes et de peu d'échantillons peut entraîner l'identification de gènes qui semblent discriminer entre les différents échantillons mais qui, en réalité, est un pur produit du hasard. Pour contourner ce problème, nous avons utilisé un «Fold-change» très élevé. Bien évidemment, ceci peut mener à la perte d'un nombre considérable de gènes ayant une importance biologique mais éviterait de retenir des faux positifs.

Lors de notre première sélection de gènes, nous avons interrogé les résultats en termes d'importance de la valeur du «Fold-change» suite à la comparaison des échantillons leucémiques aux contrôles correspondants. Un gène est sélectionné si son ratio d'expression (leucémie versus contrôle) est supérieur ou égal à 4. Cette sélection nous a permis de retenir 549 gènes dont 59 sont spécifiques aux leucémies T, 278 sont spécifiques aux leucémies B et 212 sont communs aux leucémies T et B (tableau 3.3). En notant que le nombre de gènes différentiellement exprimés est beaucoup plus important en comparant les leucémies B au CB qu'en comparant les leucémies T au CT, nous avons voulu nous assurer de la validité de ce résultat. A cette fin, nous avons utilisé le diagramme de dispersion du programme MAS5. Le même résultat a été observé (figure 3.11. C). Une des explications possibles serait la présence d'une différence au niveau de la capacité de prolifération des cellules B normales et des

cellules B leucémiques. En effet, les cellules provenant des 3 sous-types de leucémies B triées ont été extraites à partir des ganglions de souris malades. Lorsque ces souris atteignent le stade terminal, les lymphocytes B (L.B) anormaux quittent le sang et la moelle osseuse pour aller s'accumuler au niveau des ganglions où ils prolifèrent d'une façon anarchique (expliquant l'augmentation de la taille de ce réservoir). Par contre, les cellules du CB provenant des souris normales ont été extraites à partir de la rate où migrent les LB suite à leur maturation au niveau de la moelle osseuse. Au niveau de la rate, les LB matures sont au repos (état de quiescence) en l'attente d'une stimulation antigénique (Cordelier D. , 1975). Quant aux cellules provenant des leucémies T, elles ont été extraites à partir du thymus considéré comme étant un réservoir et un endroit favorable à la prolifération de

ces dernières. Les cellules provenant du CT et qui sont $CD4^+/CD8^+$ (immatures) ont aussi été extraites à partir du thymus. La migration des cellules souches lymphoïdes de la moelle osseuse vers le thymus est un phénomène continu. La différenciation intra-thymique des cellules de la lignée T est associée à une prolifération active induite par l'*IL-1* et l'*IL-7* produites par les cellules stromales thymiques (Cordelier D. , 1975). Pour résumer ce qui précède, l'explication possible quant au nombre plus élevé de gènes différenciellement exprimés entre les leucémies B et le CB est que les cellules du CB sont plus matures et prolifèrent beaucoup moins que celles des 3 leucémies B. Par contre, les cellules du CT sont aussi immatures que celles des 3 leucémies T qui, toutes les deux, sont en état de prolifération intense. C'est ainsi que nous pouvons expliquer l'apparition d'un plus grand nombre de gènes différenciellement exprimés entre les leucémies B et le CB qu'entre les leucémies T et le CT.

Lors de cette analyse, pour sélectionner ces 549 gènes, nous avons décidé d'étudier uniquement ceux qui variaient de la même façon entre les leucémies par rapport à leur contrôle. En d'autres termes, pour sélectionner les gènes spécifiques aux leucémies T, nous ne nous sommes pas intéressées à ceux qui sont, par exemple, sur exprimés dans la leucémie T1 et sous-exprimés dans les leucémies T2 et T3 par rapport au CT. Par contre, nous nous sommes intéressés à ceux qui sont soit sur-exprimés dans les 3 leucémies T par rapport au CT soit sous-exprimés dans les 3 leucémies T par rapport au CT. Nous avons appliqué le même raisonnement pour sélectionner les gènes qui sont spécifiques aux leucémies B ou bien ceux qui sont communs aux leucémies T et B. Ceci avait pour but d'éliminer le plus possible,

les gènes qui sont différentiellement exprimés entre les leucémies par rapport à leur contrôle et qui en fait peuvent être reliés à l'hématopoïèse et à un stade de différenciation déterminé et non au mécanisme de transformation cellulaire malgré que le stade de différenciation peut aussi faire partie du processus de la leucémogénèse. Un exemple que nous pouvons citer de gènes dont l'expression est reliée au stade de différenciation est celui des gènes *LYL* et *HOX11*. Le premier est uniquement exprimé au niveau des leucémies CD4⁺/CD8⁻ (stade de thymocytes précoces). Par contre, le second est uniquement exprimé au niveau des leucémies CD4⁺/CD8⁺ (stade de thymocytes immatures) (Ferrando et al. , 2002).

Afin de visualiser les différentes catégories de gènes présentes au niveau de l'ensemble des 549 gènes, nous avons utilisé un algorithme appelé « hierarchical clustering ». Cet algorithme a aussi la capacité de regrouper les échantillons en se basant sur la similitude du patron d'expression de leurs gènes (Figure 3.8) (Eisen et al. , 1998).

Etant donné qu'il nous était aussi possible d'étudier les variations géniques intra-leucémiques et malgré que cela différait de l'objectif de ce projet, nous avons quand même déterminé les gènes qui étaient spécifiques à chaque sous-type de leucémie (T1, T2, T3, B1, B2 et B3). Pour ceci, nous avons utilisé le diagramme de dispersion du programme MAS5 (figure 3.11. A) ainsi que le programme RMA (tableau 3.5). Nous avons ainsi pu remarquer, pour les 3 leucémies B, que la leucémie B3 contenait le moins de gènes différentiellement exprimés par rapport au CB alors que pour les 3 leucémies T, la leucémie T1 renfermait le moins de gènes différentiellement exprimés par rapport au CT. Pour ces dernières, ceci s'explique par le fait que le CT est au même stade de différenciation que la leucémie T1 (immature : CD4⁺/CD8⁺) alors qu'il diffère de celui des leucémies matures T2 (CD4⁺/CD8⁺) et T3 (CD4⁺/CD8⁻).

Suite à l'analyse par le programme RMA, nous avons remarqué la présence de 4 gènes (*RAG2*, *Vpreb1*, *Igll1* et *Enpep*) fortement exprimés dans les 3 leucémies B par rapport au CB. Compte tenu de la littérature, la forte expression de ces gènes dans nos leucémies nous permettait de les classer au stade de différenciation pré-B. En effet, *Rag2* est décrit

comme un gène dont l'expression est exclusive aux cellules lymphoïdes. Son rôle est d'activer la recombinaison V-(D)-J. Bories et al. , ont montré que *Rag2* était exprimé au stade pré-B et non au niveau des cellules B matures (Bories et al. , 1991). Le gène *Vpreb1* fait partie du complexe de la chaîne μ^+Kappa^+ des cellules pré-B (Ohnishi et Takemori, 1994). Cependant l'équipe qui l'a découvert n'a trouvé aucune expression au niveau des chaînes μ^+Kappa^+ d'une lignée de cellules B matures. Le gène *Igll1* est exprimé au niveau des lymphocytes pré-B humains en association avec le gène *Vpre-B* ainsi que la chaîne légère μ de l'immunoglobuline des récepteurs des cellules pré-B. Ce gène est exprimé d'une façon strictement dépendante de la lignée et du stade de maturation (Donohoe et Blomberg, 1997). Finalement, le gène *Enpep* est une glutamyl aminopeptidase qui a été uniquement trouvée au niveau des cellules B précoces de souris ainsi qu'au niveau de certaines cellules stromales, au niveau de la moelle osseuse, du

thymus, des cellules capillaires endothéliales du placenta et des cellules épithéliales de l'intestin et des tubules rénaux primaires. Cependant, au niveau des leucémies lymphoïdes, le transcrit de ce gène, a été trouvé restreint aux leucémies pré-B (Li et al. , 1993). Grâce à l'ensemble de ces gènes nous pouvons donc penser que nos 3 leucémies B sont au stade pré-B.

Suite à notre première sélection, nous avons exploré la littérature dans le but de trouver quelques gènes qui ont déjà été impliqués dans les leucémies lymphoïdes. Ces gènes pourraient apporter une assurance supplémentaire quant à la validité de nos résultats. Dans le tableau 3.3, nous avons cité quelques exemples de gènes dont 7 sont différentiellement exprimés entre les leucémies B et le CB et dont 5 sont différentiellement exprimés entre les leucémies T et le CT et dont le rôle dans les leucémies lymphoïdes a déjà été prouvé par d'autres équipes. C'est ainsi que notre étude nous a permis de valider l'absence d'expression du gène *Ccr1*. En effet, Trentin et al. , n'ont pu détecter aucune expression de ce gène chez les patients souffrants de MCL (mantle cell lymphoma), MZL (marginal zone B-cell lymphoma), SLL (small lymphocytic lymphoma), bien que ce gène soit présent chez 70% des patients atteints de CLL (chronic lymphocytic leukemia) et chez 40% de patients souffrants de HCL (hairy cell leukemia) (Trentin et al. , 2004).

Cr2 encore appelé *CD21*, est le récepteur des protéines du complément (C3d, C3g et iC3b) et de l'interféron. C'est une glycoprotéine de 145-kD exprimée au niveau des lymphocytes B et T ainsi qu'au niveau des cellules dendritiques folliculaires et au niveau des cellules épithéliales pharyngiales. Nous avons trouvé que ce gène était faiblement exprimé dans les leucémies de type B par rapport au CB (moins de 11 fois) et reste inchangé entre les leucémies T et le CT (tableau 3.3. A). Ce résultat est renforcé par le fait que l'expression de ce récepteur au niveau de la lignée cellulaire humaine SMS-SB (human pre-B cell-like acute lymphocytic leukemia cell line) a été décrite comme inexistante (White et al. , 1997).

Stmn1 encore appelée *Op18* est une oncoprotéine. *Stmn1* est une phosphoprotéine cytosolique majeure faisant partie des cellules leucémiques (Melhem et al. , 1997). Sun et al. , ont montré que l'expression de ce gène est très élevée chez les patients atteints de leucémies aiguës lymphoblastiques de type B (B-ALL) (Sun et al. , 2004) confirmant ainsi notre résultat où nous avons trouvé que ce gène est sur-exprimé dans les leucémies B par rapport au contrôle B (plus de 6 fois) et reste inchangé entre les leucémies T et le CT.

Etv5 (ets variant gene 5) ou *ERM* (Ets-related molecule PEA3-like) appartient à la superfamille PEA3 des facteurs de transcription Ets. Ce gène est impliqué dans la tumorigénèse et interagit avec le promoteur de plusieurs gènes à travers son domaine ETS. Nous avons observé une forte expression de ce gène dans les leucémies B par rapport au CB (plus de 11 fois) » Cependant, aucune variation n'a été observée entre les leucémies T et le CT. La sur-expression de ce gène au niveau des B-CLL («B chronic lymphocytic leukaemia») a également été citée par Korz C. et al. , suggérant son implication dans la pathogenèse de cette maladie (Korz et al. , 2002).

Lef1 (lymphoid enhancer binding factor 1) est un facteur de transcription impliqué dans le développement. En plus de jouer un rôle dans plusieurs types de cancers notamment le cancer du colon, il a été découvert par Klein t al. , (Klein et al. , 2001) puis par une autre équipe comme étant sur-exprimé au niveau des B-CLL (Jelinek et al. , 2003). Nous avons aussi trouvé que ce gène est sur-exprimé dans les leucémies B par rapport au contrôle B (plus de 19 fois) et qu'il ne varie pas entre les leucémies T et le CT.

L'antigène nucléaire *Mki67* (antigen identified by monoclonal antibody Ki 67) encore noté *Ki67* est une protéine pouvant être détectée au niveau de la phase G1, S, G2 et M du cycle cellulaire mais pas au niveau de la phase G0. C'est ainsi qu'il a largement été accepté comme un marqueur de prolifération des tumeurs humaines (Diop et al. , 2005). Grâce aux résultats apportés par Niewiadomska et al. , nous avons pu confirmer sa forte expression au niveau des 3 leucémies de type B par rapport au CB (48 fois). En effet, cette équipe a montré que *Mki67* est fortement exprimé au niveau des patients atteints de B-CLL et dont le stade de la maladie est avancé (Niewiadomska et al. , 2000).

La comparaison de nos résultats avec ceux de la littérature nous a aussi permis de valider le taux élevé d'expression du gène *Igfbp4* au niveau des 3 leucémies B par rapport au CB (plus de 36 fois). En plus d'être lié à d'autres types de cancers, ce dernier a été identifié par Klein et al. , (Klein et al. , 2001) comme étant sur-exprimé au niveau des leucémies B-CLL en comparaison à des cellules normales de type B (mémoires).

Dpp4 encore appelé *CD26* est une molécule se trouvant à la surface des cellules T humaines activées jouant un rôle critique dans la régulation de la lymphopoïèse (Carbone et al. , 1995). Ce gène a été montré fortement exprimé au niveau des leucémies lymphoïdes de type T, au niveau des lymphomes (Aldinucci et al. , 2004) et au niveau des leucémies B-CLL (Bauvois et al. , 1999). Ceci confirme notre résultat où nous l'avons trouvé sur-exprimé dans les 3 leucémies B par rapport au CB (plus de 22 fois) et sur-exprimé dans les 3 leucémies T par rapport au CT. Ceci montre bien que ce gène est commun aux leucémies lymphoïdes de type T et B.

Ccr4 est impliqué dans la migration des cellules T au niveau de la peau. Ce gène est sélectivement exprimé par les cellules Th2 et par les cellules T régulatrices (Yoshie et al. , 2005). Le niveau d'expression de ce gène a été étudié par plusieurs chercheurs qui ont noté sa forte expression au niveau des leucémies ATL (Adult T-cell leukemia) (Yoshie et al. , 2005). Notre étude a aussi montré que ce gène est sur-exprimé dans les 3 leucémies T par rapport au contrôle T (plus de 2 fois). Suite à leur étude, Maecker et al.. (1996) ont montré la présence du produit du gène *Tcra* (T-cell receptor alpha chain) au niveau des leucémies lymphoïdes de type T.

Notch3 (Notch gene homolog 3 (*Drosophila*)) appartient à la famille Notch. Récemment, il a été montré impliqué dans le développement des cellules T. Son activation constitutive a été reliée au développement des leucémies lymphoïdes de type T chez les animaux et chez les humains (Campese et al. , 2003). Cependant, dans la présente étude, la forte expression de ce gène par rapport au CT, n'a été détectée qu'au niveau des leucémies T de sous-type $CD4^+/CD8^+$ (environ de 6 fois) et $CD8^+/CD4^-$ (environ de 10 fois).

Comme le gène *Notch3*, *Notch1* (Notch gene homolog 1 (*Drosophila*)) fait partie de la famille Notch. Ce gène est aussi impliqué dans le développement des lymphocytes T (Liu et al. , 2003). En comparant nos résultats avec ce qui a été publié dans la littérature (tableau 3.3), nous avons aussi trouvé que le gène *Notch1* est fortement exprimé au niveau des leucémies T ($CD4^+/CD8^+$) (environ 5 fois). Nos résultats montrent que non seulement ce gène se trouve fortement exprimé au niveau de ce même sous-type de leucémie ($CD4^+CD8^+$), mais aussi au niveau de la leucémie $CD4^+/CD8^+$ (environ de 7 fois). Mais ce qui est intéressant c'est que nous l'avons trouvé sous-exprimé au niveau de la leucémie $CD4^+/CD8^-$ (environ de 4 fois). Nous pensons qu'au niveau des leucémies T1 et T2, *Notch1* peut avoir été ciblé par une intégration virale directe (juste en amont ou en aval du gène *Notch1*) ou indirecte (en amont d'un autre gène dont le produit va déréguler le niveau d'expression de *Notch1*) qui se fait totalement au hasard d'une leucémie à une autre. Par contre, ce phénomène ne s'est pas produit au niveau de la leucémie T3 faisant en sorte que le niveau d'expression de *Notch1* est resté basal (normal). Nous pouvons aussi penser que, comme le gène *GATA-1*, *Notch1* peut être exprimé d'une façon dépendante du stade de différenciation $CD8^+$.

Malgré l'existence d'études déjà publiées concernant les leucémies lymphoïdes de type T et B, la technique des micropuces à ADN nous a permis d'identifier un nombre important de nouveaux gènes qui, potentiellement, permettraient de caractériser les leucémies de type T et les leucémies de type B. Ceci démontre le grand potentiel de cet outil. Dans ce sens, nous avons identifié 56 gènes qui n'ont jamais été associés aux leucémies lymphoïdes. Certains de ces gènes sont spécifiques aux leucémies B (Annexe D), d'autres sont spécifiques aux leucémies T (Annexe E) ou bien communs aux leucémies T et B (Annexe F). Parmi ces 56 gènes, 21 ont été identifiés dans d'autres types de cancers tels que les néoplasmes pancréatiques endocrines

(gène *Pkib* (Hansel et al. , 2004)), le cancer de la prostate (gène *Crisp3* (Kosari et al. , 2002)), le cancer du colon et du sein (gène *Arhgap8* (Johnstone et al. , 2004)), le cancer de l'ovaire et du sein ainsi que le carcinome au niveau des reins (gène *Serpinf1* (Beckers et al. , 2005)). Quinze gènes sont reliés à certaines fonctions physiologiques telles que la maintenance du fuseau méiotique (gène *Fmn2* (Ryley et al. , 2005)), la production des filaments de type III d'apparence nodulaire (à cause de l'intercalation de certaines protéines telles que la filensine et la phakinin) (gène *Bfsp2* (Conley et al. , 2000)), le conditionnement moteur du clignement des yeux (gène *Camk2d* (Park et al. , 2006)), le développement et la maintenance des synapses se trouvant au niveau des photo-récepteurs des souris (gène *Cabp4* (Maeda et al. , 2005)). Parmi les 56 gènes, nous avons aussi identifié des neurexines impliquées dans l'adhésion durant la synaptogenèse (gène *Nrxn1* (Rowen et al. , 2000)), nous avons aussi répertorié un gène impliqué dans l'élongation des *VLFA* («very long chain fatty acids») saturés et mono insaturés (gène *Elovl6* (Jakobsson et al. , 2006)). Pour d'autres gènes (5 gènes), aucune information n'était disponible dans le site du PubMed (*Olfml2b*, *Hmgcll1*, *Tmc7*, *Hectd2*, *Hist1h2bc*). Parmi les 56 gènes, 8 sont des RIKEN (fonction inconnue) et 2 n'ont pas encore été identifiés (sans nom). Nous avons aussi trouvé 1 gène qui a été associé avec des leucémies CML mais pas encore à des leucémies lymphoïdes (gène *Map4k5* (Shi et al. , 1999)). Finalement, nous avons trouvé 1 gène qui est commun aux leucémies T et B mais qui, dans la littérature, n'est connu qu'au niveau des leucémies B-CLL (gène *Etv5* (Korz et al. , 2002)).

Afin de valider le profil d'expression génique obtenu suite à notre analyse, nous avons examiné, par RT-PCR, le patron d'expression de 7 gènes (*Mettl1*, *NLN*, *Arntl2*, *Gpm6a*, *Neil3*, *Celsr1* et *Dirc2*). Ces derniers que nous avons sélectionnés parmi les 56 gènes, nous semblaient les plus intéressants. Notre choix était basé sur le fait que ces gènes étaient les moins connus et dont le ratio leucémies/contrôle était très élevé. Leur localisation pouvait être membranaire, cytoplasmique ou nucléaire. Ces derniers seront probablement impliqués dans la cascade de signalisation ou dans la régulation de gènes conduisant ainsi à l'apparition du phénotype cancéreux. Nous avons aussi favorisé les gènes qui avaient des fonctions biologiques reliées à la transduction des signaux cellulaires, à la transcription, à la réparation de l'ADN, ... (Annexe G). Finalement, l'ensemble de ces gènes devait être le plus spécifique aux leucémies lymphoïdes (tableau 3.6).

Etant donné que les leucémies analysées par les micropuces contiennent une quantité d'ARN très limitante, il nous était impossible de les utiliser pour valider les résultats obtenus par RT-PCR. Nous avons donc utilisé d'autres leucémies phénotypiquement identiques aux leucémies analysées par les micropuces à ADN. Cependant, ceci pourrait être la source d'apparition de petites variations géniques entre les différentes leucémies.

Mettl1, a été identifié par Wikman et al. , comme étant fortement exprimé au niveau du cancer des poumons (Wikman et al. , 2005). C'est un gène qui est encore très peu connu. Il contient un motif hautement conservé en acides aminés, typique des méthyltransférases et montre une haute similarité de séquence avec des protéines de plusieurs organismes tels que la levure, la Drosophile et la souris (Bahr et al. , 1999). Dernièrement, la protéine kinase B (PKB) ainsi que la kinase ribosomale S6 (RSK) ont été impliqués dans l'inactivation de *Mettl1* suite à la phosphorylation de son résidu sérine au niveau de la position 27 (Ser27) (Cartlidge et al. , 2005). Un autre gène que nous avons sélectionné est nommé *Nln*. Ce dernier code pour une neurolysine faisant partie de la famille des M3 metallopeptidases. Pour une variété d'espèces et de tissus, les peptidases de la famille M3 sont impliquées dans la dégradation des peptides ainsi que dans la synthèse des peptides neuronaux bioactifs (Chao et al. , 2006). Cette neurolysine possède un site actif dont le motif est sous la forme His-Glu-X-X-His (HEXXH). Ce motif représente une partie du site de liaison aux cofacteurs des métaux. Malgré qu'*in vitro*, la neurolysine a la capacité de cliver un certain nombre de neuropeptides, son rôle *in vivo* est impliqué dans le métabolisme des neurotensines qu'elle hydrolyse entre les résidus 10 et 11 créant ainsi des fragments qui sont courts (suspectés d'être inactifs) (Brown et al. , 2001). Suite à l'analyse par les micropuces à ADN, nous avons pu voir que *Metl1* et *Nln* sont deux gènes spécifiques aux leucémies T où ils y sont fortement exprimés (Tableau 3.6). En analysant le niveau d'expression de ces deux gènes par RT-PCR, nous avons pu montrer que, pour les leucémies non triées (figure 3.10. A), ces gènes étaient effectivement plus fortement exprimés au niveau des leucémies T5, T6 spécifiquement pour *Nln*. Par contre nous avons observé que le gène *Mettl1* s'exprime de la même façon entre les différents échantillons testés. Cependant, pour les leucémies triées (figure 3.11. A), nous avons observé que le niveau d'expression de ces gènes est élevé au niveau de l'ensemble des leucémies T et B. Etant donné que le CB2 est partiellement dégradé, il nous est impossible de dire si ces gènes sont différenciellement

exprimés entre ce contrôle et les leucémies B. En revanche et en se basant sur les résultats du tableau 3.6.B, nous pouvons supposer que *Mettl1* et *Nln* s'expriment de la même façon au niveau des leucémies B et du CB2. Pour vérifier ceci, nous avons appliqué un autre programme nommé MAS5 (voir annexe H). Celui-ci attribue une lettre à chaque gène au niveau de chaque échantillon (« A » =absence d'expression, « P » =Présence d'expression) ainsi qu'une valeur d'expression nommée « signal ». Nous avons trouvé que pour l'ensemble des 6 leucémies T et B ainsi que pour leur contrôle, la lettre attribuée était un « P ». Par contre, en examinant la valeur du signal, nous avons trouvé que celui des leucémies T est très élevé par rapport au CT et un peu plus élevé par rapport aux leucémies B et le CB. Ainsi et étant donné que la valeur du signal est identique entre les 3 leucémies B et le CB, nous pensons que la forte expression de ces deux gènes au niveau des leucémies B s'explique par le fait que l'expression constitutive de *Mettl1* et *Nln* est nécessaire pour le fonctionnement des cellules des souris saines. Cependant, une activation plus intense peut se produire au niveau des cellules T entraînerait le développement des leucémies de type T. Pour expliquer autrement le résultat obtenu pour ces deux gènes que ce soit pour les leucémies triées ou pour celles non triées, nous pouvons dire que le fait de vérifier leur niveau d'expression par RT-PCR non-quantitative peut apporter une certaine marge d'erreur aux résultats observés. En effet, pour la RT-PCR conventionnelle, lorsque le plateau est atteint, il devient plus difficile de distinguer le taux d'expression d'un gène au niveau des échantillons qui l'expriment plus ou moins fortement.

Le gène *Gmp6a* encore noté M6a, code pour une protéine transmembranaire appartenant à la famille des «myelin proteolipid protein» (PLP). M6a est uniquement présent au niveau des neurones. Malgré son identification datant de plus de 20 ans, la fonction biologique de ce gène est encore peu connue. Récemment, Alfonso et al. , ont montré que la sur-expression de M6a entraînait la formation des neurites primaires au niveau des neurones (les neurites sont des prolongements cytoplasmiques du corps cellulaire). Ils ont aussi montré que M6a jouait un rôle important dans la formation de cône de croissance et dans la formation de synapses dans la moelle épinière (Alfonso et al. , 2005). Suite à l'analyse obtenue par les micropuces à ADN, nous avons trouvé que le gène *Gmp6a* est spécifique aux leucémies B où il est fortement exprimé par rapport au contrôle B. Les résultats obtenus par RT-PCR confirment la spécificité ainsi que l'exclusivité de l'expression du gène *Gmp6a* au niveau des leucémies B triées malgré

que nous ne pouvons confirmer la faible expression de *Gpm6a* dans le CB2 par rapport aux leucémies B (figure 3.11. B). Par contre, au niveaux des leucémies non triées (figure 3.10. B) et en plus d'être très fortement exprimé au niveau des leucémies B, nous avons trouvé que ce gène s'exprime aussi au niveau des leucémies Mk4 et M4 malgré que cette expression n'est pas aussi élevée que celle observée pour les leucémies B4 et B6. Nous avons donc regardé les résultats de MAS5 (voir annexe H). Cependant, ce dernier ne nous a pas permis de comprendre les causes de la sur-expression de ces gènes au niveau des leucémies autre que B étant donné qu'il considèrerait que ce gène est présent « P » au niveau des leucémies B alors qu'il le considèrerait absent « A » au niveau du reste des leucémies. Étant donné qu'au niveau des leucémies T triées, nous n'avons trouvé aucune expression du gène *Gpm6a* alors qu'au niveau des leucémies non-triées, ce gène est plus ou moins exprimé, une des explications que nous proposons pour expliquer l'expression de ce gène au niveau des leucémies T et au niveau des leucémies non-lymphoïdes, est reliée au fait que la pureté des cellules leucémiques pourrait influencer la qualité de nos résultats. Dans ce sens, généralement, les leucémies non triées contiennent plusieurs populations leucémiques qui cohabitent et dont la présence peut fausser nos résultats. Pour le gène *Gpm6a*, nous avons aussi testé une deuxième paire d'amorces située au niveau du transcrit de ce gène (proposé par la banque de donnée Ensembl). Pour ce dernier, aucune amplification n'a été observée ni pour les leucémies ni pour les contrôles. La première explication serait en relation avec une mauvaise qualité de cette deuxième paire d'amorces. La seconde explication serait reliée au fait que notre séquence Affymetrix représenterait un transcrit autre que celui proposé par la banque de donnée Ensemble. Ce transcrit pourrait être nouveau et non découvert. Pour vérifier cette dernière hypothèse, il faudrait refaire une expérience de RT-PCR avec un oligo sens au niveau du transcrit proposé par Affymetrix et un oligo anti-sens situé au niveau de la séquence Affymetrix. Dans le cas où aucune amplification ne devrait être observée, nous pourrions confirmer cette deuxième hypothèse.

Arntl2 encore appelé *BMAL2*, appartient à la superfamille des bHLH-PAS (Hoffman et al. , (1991); Ikeda et al. , (2000) ; Reddy et al. , (1986); Taylor et Zhulin, (1999)). L'expression de ce gène a été montrée limitée au cerveau foetal et au foie humain (Ikeda et al. , 2000). Malgré que sa fonction physiologique reste encore mal comprise au niveau du foie adulte, certaines évidences suggèrent que, lorsqu'il est exprimé à un fort niveau, il dérèglerait le

rythme circadien impliqué dans la production de la mélatonine (Okano et al. , 2001). Certaines autres évidences montrent que les protéines appartenant à la superfamille des bHLH-PAS, dont celle produite par *Arntl2*, sont impliquées dans la régulation de la croissance cellulaire ainsi que dans la différenciation. Ce gène a été trouvé sous-exprimé au niveau du carcinome hépatocellulaire par Yeh et al. , Cette même équipe a cherché à déterminer l'impact de sa sur-expression sur la prolifération cellulaire sans succès (Yeh et al. , 2003). Grâce à la technique des micropuces à ADN, nous avons trouvé que *Arntl2* est un gène qui est spécifique aux leucémies B où il y est fortement exprimé (contrairement à sa sous-expression au niveau du carcinome hépatocellulaire). Une des explications qui pourraient être proposées est qu'au niveau d'une cellule hépatique, *Arntl2* interagirait avec certaines protéines qui sont différentes de celles avec lesquelles il interagirait au niveau d'une cellule de rate. Les premières vont inhiber le gène *Arntl2* lors du carcinome hépatocellulaire alors que les secondes vont l'activer lors d'une leucémie de type B. Pour les cellules triées (figure 3.11. B), nous avons validé, par RT-PCR, le résultat des micropuces à ADN que ce soit pour la première paire ou pour la deuxième d'amorces étant donné qu'aucune expression de ce gène n'a été détectée au niveau des leucémies T ni au niveau du CT2. Par contre, comme précédemment et d'après les résultats présentés au niveau du tableau 3.6 et aussi ceux du programme MAS5 (Annexe H), nous pouvons uniquement supposer que ce gène est moins fortement exprimé au niveau du CB2 qu'au niveau des leucémies B.

En fouillant encore plus dans la littérature, Hungs et al. , ont montré que, comme pour le gène *Arntl2*, le gène *Pla2g4a* (« hypoxia involved Pla2g4a gene ») est sur-exprimé au niveau de la rate de souris NOD. Ce gène contient un site potentiel de liaison à l'ARNT (« ARNT binding site (TGCGTG) ») indiquant qu'il pourrait être une cible directe d'*Arntl2* (Hungs et al. , 2006). Ces chercheurs pensent que l'expression circadienne de *Pla2g4a* corrèle avec celle des gènes *Arntl1* et *Arntl2* et qu'il pourrait être régulé par ces facteurs de transcription. Nous avons ainsi consulté nos résultats obtenus avec l'analyse des micropuces à ADN et nous avons trouvé que le gène *Pla2g4a* se trouvait effectivement sur-exprimé au niveau des leucémies B par rapport au CB (environ 3 fois plus pour B1, environ 6 fois et demi pour B2 et environ 5 fois pour B3) et ne variait presque pas entre les leucémies T et leur contrôle exactement comme

pour le gène *Arntl2* (résultats non montrés). Il serait donc intéressant de confirmer l'éventuelle existence d'une coopération entre ces deux gènes.

Neil3 est une ADN glycosylase qui a été récemment identifiée au niveau des cellules humaines. Ce gène joue un rôle crucial dans la réparation de l'ADN endommagé. Il catalyse la majorité des étapes enzymatique requises pour réparer les bases endommagées par les radicaux libres, les sites des bases perdues ainsi que les cassures simples brins (Inoue et al. , 2004). Chez la souris, l'ARNm du gène *Neil3* a été uniquement détecté au niveau du thymus et de la moelle osseuse alors que chez l'homme, il a été uniquement détecté au niveau du thymus. Chez ces deux espèces, le transcrit est légèrement détectable suggérant que *Neil3* est préférentiellement exprimé au niveau des tissus hématopoïétiques. Suite à leur étude, Torisu et al. , ont montré qu'il n'existe aucune relation entre le niveau d'expression de *Neil3* et le stage de maturation des cellules B (Torisu et al. , 2005). Comme il a été montré dans notre analyse, ce gène se trouve commun aux leucémies T et B et sur-exprimé par rapport à leur contrôle respectif. En vérifiant cette expression par RT-PCR, nous avons trouvé que pour les cellules non triées (figure 3.10. C), deux leucémies valident ce résultat que ce soit pour la première paire d'amorces (au niveau de la leucémie T5) que ce soit pour la deuxième (au niveau de la leucémie T6) alors que pour les leucémies B, seule la deuxième paire d'amorce le valide pour deux leucémies (B5 et B6). Cependant, à part les leucémies T qui expriment le plus fortement le gène *Neil3*, toutes les autres l'expriment à peu près au même niveau. En comparant ces résultats avec ceux du tableau 3.6. A et B, nous pouvons voir que ces résultats ne sont pas en parfaite corrélation. De même, pour les leucémies triées (figure 3.11. C) et pour les leucémies T, nous avons pu confirmer le résultat des micropuces uniquement avec la première paire d'amorces et pour une leucémie sur deux (T8), alors que pour les leucémies B, que ce soit pour la première ou pour la deuxième paire d'amorces, nous ne pouvons confirmer que la forte expression de ce gène. Les résultats présentés pour le gène *Neil3*, semblent moins évidents. Nous pouvons penser que la première paire d'amorces située au niveau des exons 5 et 6 du transcrit du gène *Neil3* (disponible sur le site Ensembl), serait exprimée au niveau des leucémies B et T (commune) mais serait plus spécifique aux leucémies B comme le montrent les résultats de la figure 3.11. C pour *Neil3.1* où T8 sur-exprime le CT2 (ça devrait être l'inverse). Par contre et d'après les résultats de la figure 3.11. C pour *Neil3.2*, l'augmentation de l'expression de ce gène au niveau du CT2

pourrait être expliquée par le fait que la deuxième paire d'amorces située au niveau de la séquence Affymétrie, correspondrait à un nouveau transcrit épissé du gène *Neil3* nommé AK031027. Or l'ADNc de ce dernier a été cloné à partir du thymus expliquant cette sur-expression au niveau du CT2.

Il est connu qu'une activation aberrante de la voie de signalisation WNT/PCP entraîne l'apparition d'une polarité tissulaire anormale qui représente l'une des caractéristiques des cellules cancéreuses. Cette expression aberrante entraîne aussi un faible pronostic pour les cancers humains à travers l'induction de l'invasion et de la métastase (Kato et al. , 2005). *Celsr1* est situé au niveau du noyau de la molécule de signalisation (« PCP core signaling molecule ») impliquée dans la polarité planaire notée « planar cell polarity » (PCP). Ce gène joue un rôle dans la régulation de la polarité cellulaire ainsi que dans l'extension convergente (Kato et al. , 2006). *Celsr1* est exprimé intensément au niveau du neuroépithélium immédiatement avant ou durant la neurulation (Curtin et al. , 2003). Grâce à l'analyse de l'expression différentielle des gènes, nous avons trouvé que *Celsr1* est commun aux leucémies T et B pour lesquelles il est sur-exprimé par rapport à leur contrôle respectif. Ce résultat a été confirmé par RT-PCR pour les leucémies T triées (figure 3.11. C) alors que pour les leucémies non triées (figure 3.10. C), le résultat était moins évident vu que la différence d'expression de ce gène n'était pas très prononcée suite à la comparaison de l'ensemble des leucémies entre elles. Cependant et d'une façon générale, nous pouvons noter que les leucémies non triées présentent une légère expression au niveau des leucémies non-lymphoïdes et une expression un peu plus importante au niveau des leucémies lymphoïdes (le tableau 3.6.B montre aussi que pour les leucémies myéloïdes (M), érythroïdes (E) et mégakaryocytaire (Mk), ce gène est faiblement exprimé par rapport à la moyenne des échantillons).

De manière intéressante, l'analyse des résultats des micropuces montre également que le gène *Wnt8b* est spécifique aux 3 leucémies T où il est faiblement exprimé (plus de 4 fois) par rapport au CT (résultat non montré). Or, comme cité ci-dessus, l'altération de la voie de signalisation WNT/PCP entraîne l'apparition du phénotype cancéreux. Par ailleurs, nous avons aussi trouvé qu'un autre gène faisant partie du noyau de la molécule de signalisation PCP, nommé *Ankrd6* (Kato et al. , 2005), est faiblement exprimé au niveau des 3 leucémies T par rapport au CT (plus de 4 fois et demi). Ceci suggère que l'apparition du phénotype cancéreux

au niveau des cellules T pourrait être causée par une altération de la polarité tissulaire et ainsi que de la voie de signalisation WNT/PCP suite à la dérégulation de l'expression de ces trois gènes (*Celsr1*, *Ankrd6* et *Wnt8b*).

Bodmer et al. , ont récemment identifié un nouveau gène nommé *Dirc2*. Ce dernier, dont l'expression est altérée chez les patients atteints du cancer rénal 2, est localisé à proximité d'un point de cassure impliqué dans le développement de cette maladie. Des prédictions informatiques de la protéine de ce gène ont montré une homologie significative avec différents membres de la superfamille des transporteurs appelés «major facilitator superfamily» (MFS) (Bodmer et al. , 2002). Le résultat des micropuces à ADN a montré que *Dirc2* se trouvait commun aux leucémies T et B où il est faiblement exprimé par rapport à leur contrôle respectif. En vérifiant ce résultat par RT-PCR pour les leucémies triées, nous avons trouvé que la deuxième paire d'amorces, représentant la séquence Affymetrix, valide le résultat des micropuces malgré la forte expression de ce gène au niveau de la leucémie T8. Par contre, pour la première paire d'amorces, située au niveau du transcrit du gène *Dirc2*, une leucémie T sur deux valide le résultat des micropuces. Pour les leucémies B, nous pouvons supposer que *Dirc2* est sous-exprimé par rapport au CB2 comme le laissent penser les résultats du tableau 3.6 ainsi que ceux de MAS5 étant donné de sa faible expression au niveau des leucémies B7 et B8 et de l'état dégradé du CB2 (figure 3.11. D). Comme mentionné ci-haut, nous avons trouvé que l'expression de *Dirc2* est assez élevée au niveau de la leucémie T8 alors qu'elle devrait être très faible ou même absente. L'une des hypothèses possibles est reliée au fait que pour les leucémies analysées par les micropuces à ADN ainsi que pour celles analysées par RT-PCR (leucémies triées T7), le rétrovirus murin Graffi a pu s'intégrer à proximité d'un gène X et aurait donc provoqué l'activation du gène X qui en retour aurait réprimé *Dirc2*. Ce rétrovirus a aussi pu s'intégrer immédiatement en amont de *Dirc2* entraînant son inactivation. Par contre pour la leucémie triée T8 qui exprime ce gène plus ou moins au même niveau que celui du CT2, le rétrovirus a pu s'intégrer à proximité d'un autre gène impliqué dans une voie de signalisation n'impliquant pas *Dirc2*. Cette voie aurait donc aussi la capacité d'induire la leucémie.

CHAPITRE V

CONCLUSIONS ET PERSPECTIVES

Ce mémoire décrit l'utilisation de la technique de cytométrie en flux et de la technique des micropuces à ADN (microarrays) qui ont été utilisées afin de discerner les gènes différentiellement exprimés au niveau des leucémies lymphoïdes de type T et B induites par le rétrovirus murin Graffi. A l'aide de la puissante technique des micropuces à ADN (microarrays), nous avons pu caractériser 3 leucémies T ($CD4^+/CD8^+$, $CD4^+/CD8^-$, $CD4^-/CD8^+$) et 3 leucémies B ($CD45R^{\text{low}}/CD19^+$, $CD45R^+/CD19^+$, $CD45R^+/CD19^+/SCA1^+$) induites par le rétrovirus murin Graffi, en analysant la totalité de leur expression génique. Suite à la comparaison du niveau d'expression des 45000 gènes et EST, disponibles sur la puce à ADN, entre les leucémies et leur contrôle respectif, nous avons trouvé un nombre significatif de gènes différentiellement exprimés. Parmi ces gènes, quelques-uns étaient spécifiques aux leucémies T ou aux leucémies B et quelques autres étaient communs à ces deux types de leucémies. Certains de ces gènes, n'ont jamais été associés ni aux cancers ni aux leucémies. D'autres ne sont connus que dans les leucémies non lymphoïdes. Nous avons aussi trouvé des gènes dont la fonction n'a pas encore été explorée ou qui ont préalablement été liés à d'autres types de cancers mais pas aux leucémies. Ceci supporte la possibilité que l'ensemble de ces gènes pourrait être directement ou indirectement lié au processus de conversion des cellules normales en cellules malignes.

Afin de valider le résultat des micropuces, nous avons analysé, par RT-PCR, le niveau d'expression de 7 gènes. Nous avons remarqué l'existence de quelques divergences entre les résultats obtenus suite à l'analyse par la technique des micropuces à ADN et ceux obtenus par RT-PCR notamment entre les leucémies triées et celles non triées. Principalement, ces quelques différences sont reliées à l'hétérogénéité des leucémies. En effet

et en raison de contraintes techniques, nous avons testé le niveau d'expression de 7 gènes au niveau de leucémies autres que celles analysées par les micropuces à ADN tout en sachant que les leucémies sont très hétérogènes même si elles sont très proches phénotypiquement. Une autre explication serait que la technique micropuces à ADN soit beaucoup plus sensible qu'une RT-PCR non-quantitative ou même semi-quantitative. L'intégration du rétrovirus murin Graffi peut aussi jouer un rôle dans cette divergence. Dans une leucémie donnée, ce rétrovirus peut s'intégrer au hasard, dérégulant ainsi l'expression d'un gène donné alors qu'au niveau d'une autre leucémie de même phénotype, le rétrovirus serait intégré ailleurs dérégulant l'expression d'un autre gène. Par ailleurs, une des explications possibles quant à la différence notée entre les leucémies triées et celles non triées est reliée au degré de pureté des cellules leucémiques. Finalement, nous pensons que de meilleurs résultats auraient pu être obtenus si le rapport leucémie versus moyenne, avait été beaucoup plus élevé que celui que nous avons utilisé pour la majorité de nos gènes. Effectivement, nous avons remarqué que le meilleur résultat a été obtenu pour les gènes *Arntl2*, *Gmp6a* étant donné que ce rapport était relativement élevé.

Ainsi nous pensons qu'il serait important de confirmer le niveau d'expression de ces gènes par RT-PCR quantitative (Real-Time). Pour ceci, nous devrions utiliser le peu qui reste de l'ARN des mêmes leucémies qui ont été analysées par les micropuces à ADN ainsi qu'un plus grand nombre d'échantillons. Pour poursuivre ce sujet et dans le cas où les résultats que nous obtiendrons confirmeront les résultats de cette étude, nous proposons de cibler cette recherche sur le gène *Neil3* qui semble intéressant ainsi que sur les 2 gènes *Arntl2* et *Gmp6a* qui ont donnés les meilleurs résultats.

Pour le gène *Neil3* et bien que nous nous sommes attendus à avoir une meilleurs validation de nos résultats par RT-PCR, nous avons décidé de retenir la candidature de ce gène étant donné que l'étude de l'expression de ce gène nous aurait permis de mettre en évidence un nouveau transcrit alternatif en plus d'un autre transcrit qui est déjà connu. Le nouveau transcrit n'a jamais été étudié. Il a été isolé à partir d'une banque de données de thymus (RIKEN) et nous le trouvons exprimé au niveau de plusieurs leucémies étudiées T et B ainsi qu'au niveau du CT2 (figure 3.10. C). Ce transcrit aurait donc un rôle important dans les cellules hématopoïétiques et probablement dans le cancer. Pour cette raison, nous

proposons d'étudier ces deux formes d'ARNm correspondants à *Neil3* au niveau des leucémies T et B et aussi au niveau des cellules T et B normales. Il serait intéressant de déterminer par RT-PCR et avec différentes combinaisons de différents oligos, si l'expression de ces deux isoformes serait de même intensité au niveau des leucémies lymphoïdes. Nous testerons aussi l'expression du nouveau transcrit dans différents tissus. L'obtention de l'ADNc des deux transcrits dans des vecteurs d'expression (soit par clonage soit en le demandant à d'autres équipes de recherche qui l'utilisent) serait utile pour réaliser différents tests fonctionnels.

Pour le gène *Arntl2*, étant donné qu'il nous a permis de confirmer les résultats des micropuces à ADN, nous proposons de vérifier sa sur-expression au niveau des leucémies B par la création d'autres combinaisons d'oligos. Ceci permettrait par la même occasion de vérifier qu'il ne s'agisse pas d'un nouveau transcrit alternatif. Il serait aussi nécessaire de se procurer l'ADNc de ce transcrit dans un vecteur d'expression pour pouvoir appliquer plusieurs tests fonctionnels. Comme précisé au niveau de la discussion, la deuxième paire d'amorces choisie pour valider l'expression d'*Arntl2* se situe au niveau de la séquence Affymétrie, qui elle-même se situe au niveau du 3'UTR du transcrit qui n'a pas encore été séquencé. Il serait donc très intéressant de séquencer le fragment de 800 pb issu de cette amplification pour caractériser le 3'UTR du transcrit correspondant à l'ARNm d'*Arntl2* et aussi d'étudier les éléments régulateurs de cette séquence. Plusieurs protéines sont connues pour interagir avec ce gène entre autre le gène *Pla2g4a* dont la sur-expression et la spécificité aux leucémies B ont été notées suite à notre analyse par les micropuces à ADN. Il serait donc nécessaire de valider l'expression de ce gène dans les leucémies B par RT-PCR et par la suite de faire des tests d'interaction entre ces deux gènes.

Le gène *Gmp6a* codant pour une protéine à quatre segments transmembranaire nous a aussi permis de confirmer les résultats des micropuces à ADN. Nous pensons donc que cette protéine membranaire serait fortement impliquée dans la signalisation cellulaire. Nous suggérons ainsi de nous procurer son ADNc présent dans un vecteur d'expression pour faire des tests fonctionnels. Comme tests fonctionnels, nous nous proposons d'abord d'étudier le pouvoir transformant de cette protéine. Dans le cas où le test de transformation est positif, nous nous proposons de faire des mutants de cette protéine (mutation du domaine extra-

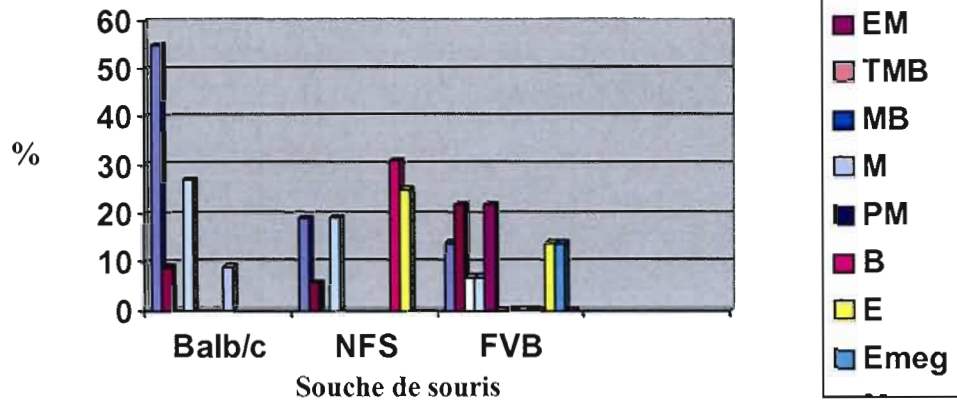
cellulaire et intra-cellulaire) et d'en voir les conséquences. Les résultats des micropuces à ADN nous ont aussi permis de découvrir un autre isoforme de ce gène noté *Gmp6b* et qui est aussi tout aussi spécifique et sur-exprimé au niveau des leucémies B. Il serait donc intéressant d'étudier comme pour le gène *Gmp6a*, son rôle dans la leucémogénèse.

D'une façon générale, pour l'ensemble de ces gènes sélectionnés ou pour d'autres gènes que nous sélectionnerons dans le cas où les résultats obtenus ne confirment pas ceux obtenus par la technique des micropuces à ADN, nous proposons d'étudier en premier lieu, leur pouvoir transformant. Il serait aussi intéressant de déterminer si l'augmentation de l'expression de certains gènes pourrait être due à une activation directe ou indirecte par le rétrovirus. Une étude *in vitro* et *in vivo* pourrait mener à une meilleure compréhension de leur fonction. Pour les gènes codant pour des protéines se liant à l'ADN, il serait nécessaire d'étudier la fonction de ces protéines et déterminer leurs partenaires d'interaction. Pour ceux qui codent pour des récepteurs membranaires ou impliqués dans la signalisation cellulaire, il serait intéressant d'étudier la voie qui pourrait être impliquée dans la transformation cellulaire. L'un des futurs objectifs de ce projet est de vérifier l'expression des gènes sélectionnés chez l'humain dans le but de trouver des gènes marqueurs pour discriminer les leucémies entre elles à des fins thérapeutiques.

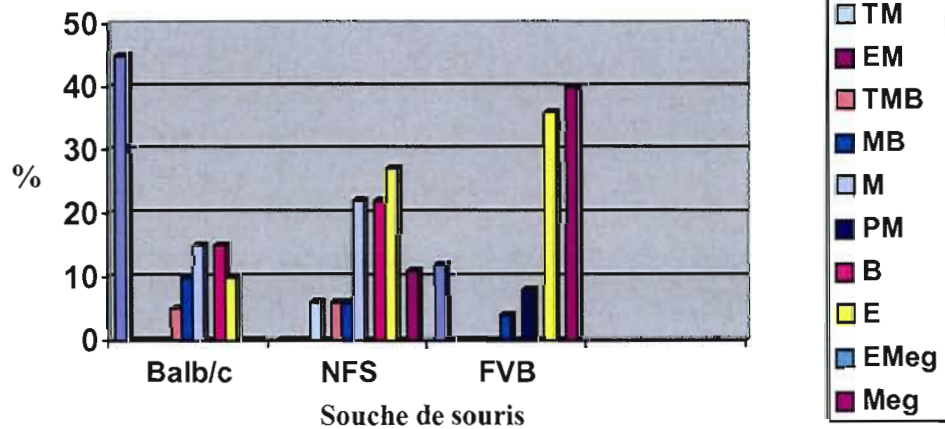
En conclusion, l'utilisation de la technique des micropuces à ADN nous a prouvé sa puissance quant à l'impressionnante quantité d'informations qu'elle a été capable de nous fournir concernant les gènes qui seraient directement ou indirectement impliqués dans leucémies lymphoïdes. Cette technique a mené à une meilleure caractérisation et à une meilleure compréhension des mécanismes qui seraient responsables de l'induction des leucémies chez le modèle murin. La poursuite de ce projet par des analyses fonctionnelles permettrait de mieux identifier les gènes les plus appropriés dans l'induction des leucémies lymphoïdes.

ANNEXE

**Pourcentage de chaque type
de leucémie induite par le clone moléculaire
GV-1.2 du rétrovirus murin Graffi**



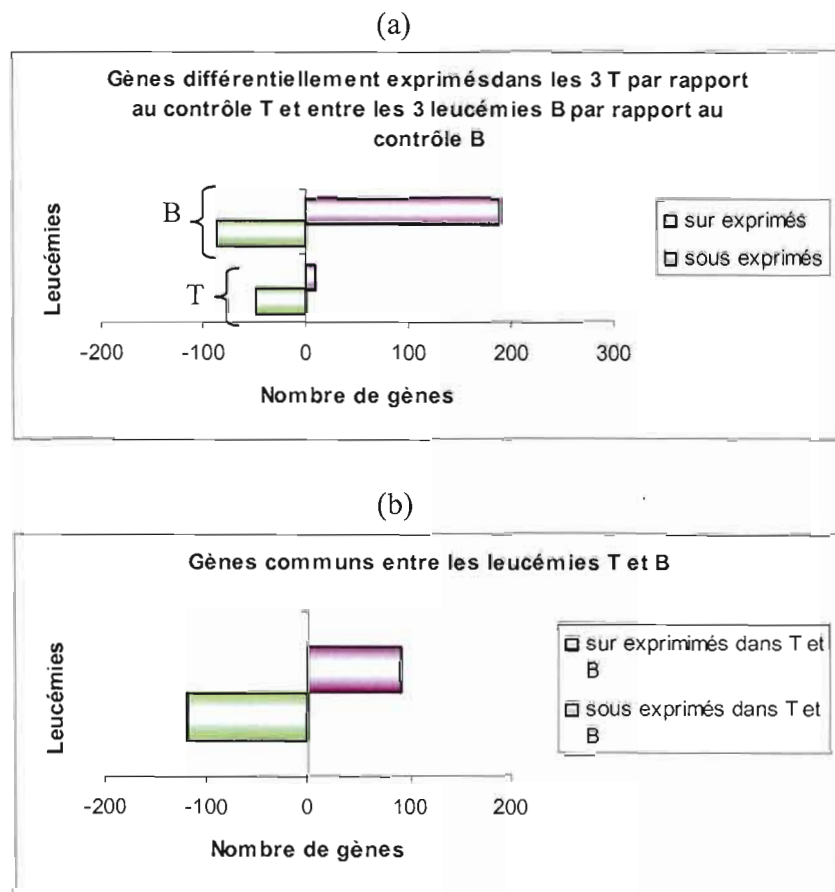
**Pourcentage de chaque type
de leucémie induite par le clone moléculaire
GV-1.4 du rétrovirus murin Graffi**



Annexe A Pathogénèse du rétrovirus murin Graffi (Voisin et al. , 2006). Légende : **T** : leucémie lymphoïde de type T; **E** : leucémie érythroïde; **M** : leucémie myéloïde; **B** : leucémie lymphoïde de type B; **P** : primitive ; **Meg** : leucémie mégakaryocytaire. Trois souches de souris ont été utilisées : Balb/c, NFS et FVB.

Souche	Numéro	Leucémie	cellules leucémiques	cellules normales	FOG
NFS1,2	VV41.1	E	97	3	3
NFS1,2	VV38.3	E	65	35	3
NFS1,2	VV42.4	E	60	40	3
NFS1,2	VV42.5	E	-	-	absent
FVB1,2	VV29.7	E	91	9	2
FVB1,2	VV29.4	E	90	10	2
FVB1,2	VV29.8	E	85	15	absent
FVB1,2	VV29.2	E	80	20	2
FVB1,2	VV12.4	E	90	10	1
Balb 1,4	VV07.1	E	75	25	3
Balb 1,4	VV34.4	E	70	30	absent
NFS1,4	VV06.8	E	60	40	2
NFS1,4	VV03.4	E	82	18	3
NFS 1,4	VV39.2	E	-	-	2
-	-	E	-	-	1
FVB1,4	VV30.4	E	91	9	3
FVB1,2	VV29.3	EMk	80	20	2
FVB1,2	VV29.5	EMk	85	15	2
NFS1,4	VV40.x	Mk	80	20	absent
NFS 1,4	VV39.3	Mk	-	-	1
FVB1,2	VV29.6	Mk	90	10	absent
FVB1,4	VV30.8	Mk	80	20	1
FVB1,4	VV30.9	Mk	53	36	1
FVB1,4	VV32.2	Mk	40	60	absent
FVB1,4	VV09.4	Mk	70	30	3
FVB1,4	VV10.4	Mk	90	10	-
FVB1,4	VV10.6	Mk	75	25	3
-	-	Mk	-	-	absent
Balbc 1,2	VV28.5	M	80	20	absent
Balb 1,4	VV33.4	M	80	20	absent
NFS1,4	VV04.2	M	80	20	2
NFS 1,4	VV39.1	M			1
-	-	M	-	-	absent
Balbc 1,2	VV25.4	T	60	40	absent
Balbc 1,2	VV25.1	T	60	40	absent
Balbc 1,2	VV28.3	T	90	10	absent
NFS1,2	VV05.4	B	58	36	absent
NFS1,2	VV42.2	B	70	30	absent
NFS1,2	VV41.4	B	60	40	1
-	-	B	-	-	-

Annexe B Caractéristiques des leucémies choisies pour tester le niveau d'expression de GATA-1, GATA-2 et de GATA-3. Les cellules normales et les cellules leucémiques ont une localisation bien particulière lorsque nous les observons sur l'écran de l'appareil du Facsvantage. Le programme de ce dernier permet de déterminer le pourcentage des cellules leucémiques ainsi que celui des cellules normales. Ces pourcentages dépendent de la taille de chacune de ces deux populations. Au niveau de la dernière colonne, nous avons indiqué les résultats obtenus suite à l'hybridation des ARN provenant de chacune de ses leucémies avec la sonde *Fog-1*. Le chiffre « 1 » indique que le gène est légèrement exprimé, le chiffre « 2 » indique que *Fog-1* est moyennement exprimé et le chiffre « 3 » indique que *Fog-1* est très fortement exprimé. Par contre, la notation « absent » veut dire que *Fog-1* n'est pas exprimé. Les tirés indiquent une absence d'information.



Annexe C Figure récapitulant le nombre de gènes qui a été trouvé dans chaque groupe de leucémies. (a) : Dans cette figure, se trouvent les gènes qui sont spécifiques aux leucémies B ainsi que ceux qui sont spécifiques aux leucémies T. Les gènes ont été considérés comme étant spécifiques aux leucémies B lorsqu'ils étaient plus ou moins exprimés par rapport au CB mais dont l'expression ne variait pas dans les leucémies T et le CB et vice versa pour les gènes spécifiques des leucémies T; (b) : dans cette figure se trouvent les gènes qui sont communs entre leucémies B et T. Au sein d'un même groupe, nous avons séparé nos gènes selon qu'il soient sur-exprimés par rapport au contrôle (colonnes colorées en rose) ou bien sous exprimés par rapport au même contrôle (colonnes colorées en vert).

ProbesetIDs	B1-CB	B2-CB	B3-CB	CT	T1	T2	T3	CB	B1	B2	B3	M	CE	E1	E2	E3	Meg1	Meg2	Meg3	Gene Title	Gene Symbol
1419219_at	-4.999	-4.84	-4.41	-1.2	-1.69	-1.188	-1.37	5.23	0.27	0.43	0.86	0.597	-0.66	-1.16	-1.41	-1.25	-1.7	-1.7	-1.7	cytochrome P450, family 4, subfamily f, polypeptide 18	Cyp4f18
1421138_a_at	-4.202	-3.62	-3.48	-0.52	-0.85	-0.844	-0.53	1.56	-0.65	-0.06	1.05	1.19	-0.3	-0.59	-0.69	-0.9	-0.87	-0.42	0.3	protein kinase inhibitor beta, cAMP dependent, testis specific	Pkinb
1449838_at	-3.103	-3.02	-3.05	-0.48	-0.75	-0.184	-0.47	2.65	-0.45	-0.37	-0.39	-0.68	-0.52	-0.37	-0.55	-0.46	2.16	-0.21	1.0	cysteine-rich secretory protein 3	Crs3
1418809_at	-3.036	-3.02	-3.83	-1.39	-1.7	-1.696	-1.51	1.56	-1.48	-1.46	-2.27	1.3	2.37	0.25	-1.09	-1.81	0.06	1.57	3.25	paired-Ig-like receptor A1	Pira1
1451463_at	-2.953	-4.15	-2.9	-0.59	-0.59	-0.883	-0.34	3.78	1.13	-0.36	0.88	0.02	-0.2	-0.95	-0.5	-0.52	-0.41	-0.16	0.0	Rho GTPase activating protein 8	Rhga8
1426959_at	2.187	2.458	2.234	1.833	1.416	1.539	1.16	-1.1	1.08	1.36	1.13	-1.65	-1.52	-0.71	-0.12	0.4	-1.33	-2.12	-1.6	3-hydroxybutyrate dehydrogenase (heart, mitochondrial)	Bdh
1429430_at	2.262	2.163	2.077	-0.29	-0.3	1.043	-0.66	0.67	2.33	2.12	2.15	-0.59	-0.54	-0.85	-1.14	-0.71	-1	-0.08	-0.6	RIKEN cDNA A030012M09 gene	A030012M09R
1450063_at	2.585	3.046	3.038	-0.43	-0.76	-0.624	-0.62	-0.4	2.14	2.6	2.58	-0.57	-0.44	-0.48	-0.76	-0.67	-0.44	-0.64	-0.4	formin 2	Fmn2
1434463_at	2.978	3.361	3.39	-1.2	-1.65	-1.675	-1.49	1.77	4.75	6.13	6.16	-1.62	-1.33	-1.15	-1.5	-1.37	-1.16	-1.52	-1.1	beaded filament structural protein 2, phakinin	Bfsp2
1419153_at	3.073	3.066	3.176	0.857	0.853	0.656	1.04	-1.9	1.13	1.37	1.24	-0.41	0.47	-0.11	-0.15	-0.18	-0.88	-2.84	-1.3	RIKEN cDNA 2810417H13 gene	2810417H13R
1439789_at	3.086	2.886	2.517	-1.43	-1.54	-1.639	-1.55	1.75	4.82	4.53	4.25	-1.75	-1.54	-1.28	-1.03	-1.35	-1.46	-0.4	-0.5	—	—
1423007_a_at	3.124	3.223	2.875	-0.85	-1.26	-0.889	-0.86	-0	3.08	3.18	2.65	-0.76	1.33	-1.03	-0.86	-0.99	-0.71	-1	-0.6	glial cell line derived neurotrophic factor family receptor alpha 2	Gfra2
1453724_a_at	3.179	2.888	3.044	-0.27	-0.53	-0.759	-0.67	-0.7	2.48	2.2	2.35	-0.07	-0.5	-0.55	-0.36	-0.53	-0.47	-1.1	-0.5	serine (or cysteine) proteinase inhibitor, clade F, member 1	Serpinf1
1429688_at	3.237	2.775	2.753	-0.37	-0.85	-0.492	-0.74	-0.4	2.85	2.4	2.37	-0.62	-0.63	-0.53	-0.64	-0.53	-0.52	-0.63	-0.7	aryl hydrocarbon receptor nuclear translocator-like 2	Ahr2
1425878_at	2.003	2.881	2.528	-0.47	-0.82	-0.48	-0.44	-0	1.98	2.85	2.5	-0.87	-0.1	-0.77	-0.68	-0.78	-0.15	-1.03	-0.7	calcium binding protein 4	Cabp4
1436948_a_at	2.165	6.278	6.556	-1.02	-0.54	-0.827	-0.32	-0.9	1.36	3.37	3.65	-1.06	-1.08	-1.02	-0.79	-0.93	-1.06	-0.87	-0.9	RIKEN cDNA 6430550H21 gene	6430550H21R
1453060_at	2.587	3.681	3.411	-0.59	-0.66	-0.714	-0.72	-0.5	2.04	3.12	2.31	-0.51	-0.34	-1.02	-0.51	-0.3	-0.97	-0.78	-0.4	RIKEN cDNA 6530413H01 gene	6530413H01R
1444031_at	2.758	3.048	2.34	-0.68	-0.18	-0.507	-0.66	-0.2	2.32	2.82	2.11	-1.07	-0.72	-0.72	-0.65	-0.65	-0.42	-0.31	-0.5	calcium/calmodulin-dependent protein kinase II, delta	Camk2d
1458396_at	3.042	2.746	6.174	-0.63	-1.12	-1.072	-0.81	-0.8	2.25	1.95	6.42	-0.03	-0.46	-0.9	-0.88	-0.76	-0.67	-0.7	-0.9	15 days embryo head cDNA, RIKEN full-length enriched library, —	—
1423915_at	3.058	3.13	2.164	-0.53	-0.64	-0.62	-0.56	-0.6	2.5	2.57	1.6	-0.41	-0.19	-0.64	-0.25	-0.36	-0.67	-0.71	-0.5	olfactomedin-like 2B	—
1435361_at	3.078	3.005	3.161	-0.48	-0.85	-0.916	-0.65	-0.5	2.55	2.48	2.25	-0.64	-0.14	-0.77	-0.3	-0.85	-0.59	-0.98	-0.6	cDNA sequence AK129018	AK129018
1426442_at	3.148	2.788	2.425	-0.47	-0.53	-0.509	-0.38	-0.5	2.65	2.29	1.93	-0.61	-0.47	-0.46	-0.53	-0.52	-0.65	-0.7	-0.5	glycoprotein m6a	Gpm6a
1448460_at	3.257	3.244	2.176	-0.37	-0.16	-0.161	-0.49	-0.8	2.47	2.46	1.07	-0.21	-0.31	-0.58	-0.74	-0.78	-0.78	-0.41	-0.6	activin A receptor, type 1	Actr1
1434913_at	3.36	2.568	4.285	-0.87	-0.97	-0.784	-0.92	-0.6	2.72	1.23	2.65	-1	-0.56	-0.42	-0.26	-0.5	-0.06	-0.8	-0.5	3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase-like 1	—
1441977_at	3.51	3.654	2.077	-1.45	-1.44	-1.396	-1.21	0.89	4.4	4.49	3.76	-1.36	-1.26	-1.12	-1.19	-1.34	-1.73	-0.56	0.52	RIKEN cDNA 9630023C09 gene	9630023C09R
1430479_at	3.628	2.889	1.83	-0.75	-0.48	-0.545	-0.37	-0.5	3.09	2.47	1.3	-0.75	-0.46	-0.61	-0.45	-0.5	-0.44	-0.37	-0.6	RIKEN cDNA 2010007H06 gene	2010007H06R
1446748_at	4.031	3.537	2.676	-1.05	-0.66	-0.848	-0.66	-0.2	3.78	3.3	2.44	-1.06	-0.92	-0.95	-0.77	-0.91	-0.67	-0.1	-0.7	RIKEN cDNA 2010007H06 gene	2010007H06R
1419152_at	4.063	4.703	4.472	1.255	0.76	1.571	1.06	-2.6	1.45	2.08	1.45	-1.01	-0.41	-0.38	-0.39	-0.53	-1.55	-2.51	-1.6	RIKEN cDNA 2810417H13 gene	2810417H13R
1454691_at	4.213	3.582	4.182	-0.94	-0.95	-1.184	-0.75	-1.3	2.94	2.29	2.31	0.594	-0.95	-0.71	-1.1	-1.16	0.16	1.05	-0.4	neurexin I	Nrxn1
1456981_at	4.988	4.68	3.078	-1.22	-1.03	-1.163	-1.38	-0.7	4.21	3.77	2.38	2.043	-0.25	-1.06	-0.9	-1.2	-1.19	-1.02	-1.3	transmembrane channel-like gene family 7	—
1448529_at	5.316	3.888	4.285	-1.41	-1.31	-1.458	-1.32	-1.1	4.22	2.73	3.13	2.007	-0.4	-1.02	-0.31	-0.56	-1.13	-1.07	-1	thrombospondin	Thsd
1437856_at	5.385	4.853	4.482	-1.21	-1.04	-0.946	-0.89	-2.2	3.15	2.82	2.26	-0.49	-0.11	-0.26	0.307	0.41	-0.72	-0.6	-0.5	inositol polyphosphate multikinase	Ipmk
1425942_a_at	6.929	6.346	2.55	-0.72	-1.16	-1.267	-0.82	-1	5.88	5.3	1.5	-1.28	-0.87	-0.3	-1.27	-1.23	-0.86	-1.22	-0.6	glycoprotein m6b	Gpm6b

Annexe D Gènes spécifiques aux leucémies B. Suite à la première sélection et à la recherche bibliographique, 31 gènes, dont le rôle dans la leucémie reste encore inconnu, ont été retenus. **Vert** : gènes sous exprimés par rapport au CB; **Rouge** : gènes sur exprimés par rapport au CB; **jaune** : gènes ayant un rôle dans les fonctions physiologiques; **Rose** : gènes connus dans d'autres types de cancer; **bleu blanc** : aucune information disponible, fonction inconnue, blanc : gènes RIKEN. Les valeurs sont exprimées en logarithme base 2. Ainsi une valeur de 4 signifie une variation d'expression de 2⁴ et donc de 16 fois.

Proteins IDs	T1-CT	T2-CT	T3-CT	Control	T1	T2	T3	CB	B1	B2	B3	IM	OE	E1	E2	E3	Meg1	Meg2	Meg3	Gene Title	Gene Symbol
1405702_a.at	-2.792	-3.33	-5.75	5.072	2.391	1.833	-0.57	-0.7	-0.64	-0.92	-0.96	-0.68	-0.87	-0.9	-0.69	-0.79	-0.78	-0.23	-0.3	ORF-related orphan receptor gamma	Rorc
1421919_a.at	-2.679	-6.65	-6.65	3.078	3.295	-0.672	-0.67	-0.6	-0.6	-0.59	-0.93	-0.75	-0.55	-0.83	-0.5	-0.47	-0.45	-0.73	-0.8	chemokine (C-C motif) receptor 9	Ccr9
1424015_a.at	-3.268	-3.48	-3.78	-0.38	-3.67	-3.854	-4.16	0.34	0.57	0.28	0.73	0.59	0.97	0.35	0.792	0.59	0.79	1.57	1.3	Ras6 interacting protein 1	Rasip1
1448321_a.at	-3.039	-3.03	-2.95	2.775	-0.38	-0.315	-0.24	-0.2	-0.34	-0.4	-0.28	-0.25	0.34	-0.39	-0.24	-0.22	0.55	-0.51	-0.2	SPARC related modular calcium binding 1	Sparc1
1451555_a.at	2.395	2.026	2.87	-1.53	0.795	1.727	0.34	-0.3	-0.35	-0.21	-0.2	-0.1	-1.04	-0.08	-0.3	0.31	0.57	-0.03	0.5	neurolysin (metallopeptidase M3 family)	Nln
1419122_a.at	2.353	2.762	2.845	-1.57	0.764	0.535	1.26	0.91	-0.4	-0.06	-0.14	0.93	-0.9	-0.11	0.76	0.81	0.14	-0.01	-0.1	methyltransferase-like 1	Mettl1
1460368_a.at	3.386	2.644	5.363	-0.22	3.153	2.424	5.73	-1.1	-1.15	-1.25	-1.47	-1.17	-0.88	-1.28	-1.31	-1.25	-0.47	-1.12	1.2	membrane protein, palmitoylated 4 (MAGUK p65 subfamily member)	Mags4
1433944_a.at	3.588	3.277	2.185	-0.5	3.065	2.777	1.65	-0.9	-1.06	-1	-0.96	0.24	-0.83	-0.86	-1.01	-0.9	-0.49	0.7	-0.7	HECT domain containing 2	Hect2
1427600_a.at	7.278	7.284	6.979	-1.23	6.948	6.853	9.78	-1.7	-1.89	-1.86	-1.8	-1.93	-1.71	-1.6	-1.91	-1.72	2.47	-2.07	-0.9	RIKEN cDNA A130054M15 gene	A130054M15R

Annexe E Gènes spécifiques aux leucémies T. Suite à la première sélection et à la recherche bibliographique, 8 gènes, dont le rôle dans la leucémie reste encore inconnu, ont été retenus. **Vert** : gènes sous exprimés par rapport au CT; **Rouge** : gènes sur exprimés par rapport au CB; **Jaune** : gènes ayant un rôle dans les fonctions physiologiques; **Rose** : gènes connus dans d'autres types de cancer; **bleu-rose** : aucune information disponible, fonction inconnue, blanc : gènes RIKEN. Les valeurs sont exprimées en logarithme base 2. Ainsi une valeur de 4 signifie une variation d'expression de 2^4 et donc de 16 fois.

Probes IDs	T1-CT	T2-CT	T3-CT	B1-CB	B2-CB	B3	M	OE	E1	E2	E3	IM1	IM2	IM3	Gene Title	Gene Symbol	
1418172_at	-3.786	-3.69	-3.41	-2.86	-2.6	-1.62	-2.9	-2.74	-2.46	-2.55	-2.51	-2.45	-1.01	-0.76	hemie binding protein 1	Hebp1	
145454_at	-3.255	-4.03	-4.45	-1.13	-1.72	-4.71	-3.06	-0.9	-1.96	-1.42	-0.65	-1.1	-1.56	-0.73	disrupted in renal carcinoma 2 (human)	Drc2	
1418072_at	-3.110	-2.43	-2.42	-4.15	-3.02	-3.495	-2.6	-1.09	-1.65	-2.28	-1.78	-1.44	-1.12	-0.71	histone 1, H2bc	Hist1h2bc	
1430523_s_at	-2.784	-2.69	-1.79	-6.67	-5.10	-6.163	-1.5	-1.74	-1.64	-0.74	-0.55	-0.31	-0.67	-0.3	immunoglobulin lambda chain, variable 1	Ighl	
1427709_x_at	-2.102	-1.72	-1.22	-5.33	-5.30	-5.954	-1.7	-1.5	-1.15	-0.65	-0.42	-0.21	-0.26	-0.64	immunoglobulin kappa chain variable 21 (V21)	Ighk21	
1427376_at	-2.04	-2.41	-2.25	-1.54	-1.22	-1.323	-1.22	-1.32	-1.39	-1.43	-1.53	-1.01	-0.69	-0.79	mitogen-activated protein kinase kinase kinase 5	Mkk45	
1418243_at	-1.133	-1.57	-1.42	-4.2	-4.04	-3.607	-1.3	-2.4	-2.87	-2.72	-1.55	-2.64	-2.48	-2.05	fibron A	Fcna	
1429491_s_at	1.421	1.453	1.555	2.443	2.242	2.652	0.96	0.6	0.22	0.24	-2.51	-0.05	0.71	0.33	Rap1 interacting factor 1 homolog (yeast)	Rif1	
1430465_at	1.58	2.657	1.08	2.763	3.025	2.706	-0.74	1.34	1.32	0.1	-2.05	0.705	0.055	-0.32	RIKEN cDNA 301009J007 gene	301009J007R0	
1430111_at	1.621	2.42	3.262	1.251	2.705	2.556	-0.92	0.76	1.5	1.3	-0.59	2.091	0.72	-1.02	branched chain aminotransferase 1, cytosolic	Bcat1	
1417404_at	1.838	1.044	2.95	2.043	1.757	1.48	-1.27	0.54	-0.23	1.7	-3.45	3.05	3.72	-1	ELOVL family member 6, elongation of long chain fatty acids (yeast)	Elo6	
1447102_at	1.553	1.670	1.213	1.26	2.504	1.581	-1.02	0.52	1.55	0.5	-1.38	0.04	0.778	0.13	RAV1-like 1 (S. cerevisiae)	Rav1l1	
1450071_at	2.2	2.67	3.84	2.555	1.70	3.452	-0.62	1.51	2.33	1.5	-0.97	1.721	3.04	1.59	-	-	-
1418925_at	2.306	1.565	1.06	3.997	1.653	2.445	1.73	1.65	2.26	1.9	-1.27	1.725	1.4	-1.01	calcium EGF-like seven-pass G-type receptor 1	Cesr1	
1453753_at	2.472	3.648	2.011	3.548	1.677	3.561	-1.37	1.1	2.26	1.4	-2.36	1.003	1.573	2.2	RIKEN cDNA 2810047.02 gene	2810047.02R0	
1428142_at	2.658	2.207	3.471	4.04	1.574	3.66	-1.05	0.71	0.9	2.3	-0.26	3.702	3.704	3.4	ets variant gene 5	Etv5	
1440633_at	2.874	3.445	1.882	3.553	3.335	3.107	-2	0.80	1.4	-0.31	-2.45	1.02	1.02	-0.41	nei like 3 (E. coli)	Nei3	

Annexe F Gènes communs aux leucémies T et B. Suite à la première sélection et à la recherche bibliographique, 17 gènes, dont le rôle dans la leucémie reste encore inconnu, ont été retenus. **Vert** : gènes sous-exprimés dans leucémies T par rapport au CT et dans les leucémies B par rapport au CB; **Rouge** : gènes sur exprimés par rapport au CB; **Jaune** : gènes ayant un rôle dans les fonctions physiologiques; **Rose** : gènes connus dans d'autres types de cancer; **bleu foncé** : aucune information disponible, fonction inconnue; **Aubergine** : connus au niveau des leucémies non lymphoïdes et non au niveau des leucémies lymphoïdes; **Gris** : connus uniquement dans l'un des deux types de leucémies T ou B; blanc : gènes RIKEN.

	Gene Title	Gene Symbol	GO Biological Process Description	GO Molecular Function Description	GO Cellular Component Description
1419122_at	methyltransferase-like 1	Mettl1	transport tRNA processing tRNA modification	transporter activity methyltransferase activity tRNA (guanine-N7-)-methyltransferase S-adenosylmethionine-dependent transferase activity	membrane nucleus
1424981_at	neurolysin (metallopeptidase M3 family)	Nln	proteolysis	metalloendopeptidase activity peptidase activity zinc ion binding hydrolase activity metallopeptidase activity	
1429688_at	aryl hydrocarbon receptor nuclear translocator-like 2	Arntl2	regulation of transcription from RNA polymerase II promoter signal transduction circadian rhythm regulation of transcription, DNA-dependent regulation of transcription	transcription factor activity Signal transducer activity Transcription regulator activity	nucleus
1456741_s_at	glycoprotein m6a	Gpm6a			integral to membrane membrane
1440633_at	nei like 3 (E. coli)	Nei3	DNA repair response to DNA damage base-excision repair	DNA binding zinc ion binding purine-specific oxidized base lesion DNA N-glycosylase activity metal ion binding damaged DNA binding	nucleus
1418925_at	cadherin EGF LAG seven-pass G-type receptor 1	Celsr1	homophilic cell adhesion signal transduction G-protein coupled receptor protein signaling pathway neuropeptide signaling pathway development locomotory behavior establishment of polarity of embryonic epithelium inner ear morphogenesis	signal transducer activity receptor activity G-protein coupled receptor activity structural molecule activity calcium ion binding brain-specific angiogenesis inhibitor activity	extracellular space membrane integral to membrane
1454654_at	disrupted in renal carcinoma 2 (human) similar to Disrupted in renal carcinoma 2	Dirc2 LOC547255		peptidase activity	

Annexe G Fonctions biologiques des 7 gènes sélectionnés pour confirmer le résultat des micropuces à ADN. En plus de présenter les fonctions biologiques de chacun des 7 gènes, dans ce tableau, nous pouvons aussi trouver leurs fonctions moléculaires ainsi que leur localisation cellulaire. Ce tableau est extrait à partir du site Affymetrix (www.affymetrix.com) dans lequel se trouvent toutes les données concernant l'ensemble des gènes présents sur notre puce à ADN.

Probeset	Gene	CT	T1	T2	T3	C8	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
1419122_at	Meth1	P(13.1)	P(139.4)	P(101.6)	P(125.9)	P(38.3)	P(21.2)	P(63.6)	P(39.4)	P(50.7)	P(38.9)	P(70.9)	P(40.7)	P(25.9)	P(54.7)	P(41.3)
1424081_at	Nln	P(15.3)	P(139.8)	P(203.2)	P(65.9)	P(43.7)	P(37)	P(42.6)	P(32.9)	P(64.6)	P(51.5)	P(84)	P(83.3)	P(23.8)	P(75.4)	P(33.3)
1429688_at	Amli2	A(8.6)	A(0.1)	A(0.4)	A(1.5)	A(4.6)	P(60.5)	P(67.3)	P(54.1)	A(1.1)	A(0.3)	A(1.6)	A(0.4)	A(0.9)	A(0.9)	A(0.4)
1456741_s_at	Gmpb6a	A(4.8)	A(6.1)	A(3.6)	A(3.5)	P(14.1)	P(145.3)	P(167.2)	P(81.2)	A(3)	A(2.6)	A(5.1)	A(8.5)	A(0.3)	A(0.3)	A(8.7)
1440633_at	Neil3	A(10.8)	P(125.2)	P(167.6)	P(38)	A(2.8)	P(30.5)	P(176.4)	P(71.4)	P(47.5)	P(106.1)	P(20.5)	P(54.3)	P(14.8)	P(28.1)	P(27.9)
1418825_at	Celsr1	P(40.8)	P(275.2)	P(140.5)	P(117.2)	A(2.9)	P(78.3)	P(84.9)	P(53.5)	A(12.3)	A(9.7)	A(0.4)	A(14.3)	A(0.4)	A(2.4)	A(5.1)
1454654_at	Dir2	P(12.1)	P(139.4)	P(101.6)	P(125.9)	P(38.3)	P(21.2)	P(63.6)	P(39.4)	P(33.3)	P(11.1)	P(75.7)	P(27.9)	P(13.6)	P(91.8)	P(41.5)

Annexe H Résultats obtenus par MAS5 suite à l'analyse de l'expression des 7 gènes.

En se basant sur l'intensité d'expression d'un gène, le programme MAS5 lui attribue la lettre A, M ou P. Lorsque l'intensité de la fluorescence émise par le gène est nulle, ce programme lui attribue la lettre « A » (« A » pour absent). Par contre, la lettre « P » est attribuée si le gène est exprimé (« P » pour présent). Si cette expression est marginale, la lettre « M » est attribuée. En vert se trouvent les gènes qui sont absents et en rouge se trouvent ceux qui sont présents. Les chiffres entre parenthèses représentent la valeur attribuée à chaque gène. C'est selon cette valeur qu'un gène est désigné par telle ou telle lettre.

BIBLIOGRAPHIE

- ALDINUCCI D, POLETTO D, LORENZON D, NANNI P, DEGAN M, OLIVO K, RAPANA B, PINTO A, GATTEI V. 2004. CD26 expression correlates with a reduced sensitivity to 2'-deoxycoformycin-induced growth inhibition and apoptosis in T-cell leukemia/lymphomas. *Clin Cancer Res.* 15;10(2):508-20.
- ALFONSO J, FERNANDEZ ME, COOPER B, FLUGGE G, FRASCH AC. 2005. The stress-regulated protein M6a is a key modulator for neurite outgrowth and filopodium/spine formation. *Proc Natl Acad Sci U S A.* 22;102(47):17196-201. Epub.
- ALLEN TD, DEXTER TM. 1995. The biology of haematopoiesis. In *Oxford Textbook of Oncology* (ed. M Peckham, H Pinedo and U Veronesi). 1593-1600.
- BAHR A, HANKELN T, FIEDLER T, HEGEMANN J, SCHMIDT ER. 1999. Molecular analysis of METTL1, a novel human methyltransferase-like gene with a high degree of phylogenetic conservation. *Genomics.* 57(3):424-428.
- BAUVOIS B, DE MEESTER I, DUMONT J, ROUILLARD D, ZHAO HX, BOSMANS E. 1999. Constitutive expression of CD26/dipeptidylpeptidase IV on peripheral blood B lymphocytes of patients with B chronic lymphocytic leukaemia. *Br J Cancer.* 79 (7-8):1042-8.
- BECKERS J, HERRMANN F, RIEGER S, DROBYSHEV AL, HORSCH M, HRABE DE ANGELIS M, SELIGER B. 2005. Identification and validation of novel ERBB2 (HER2, NEU) targets including genes involved in angiogenesis. *Int J Cancer.* 114(4):590-7.
- BERNARDI R, GRISENDI S, PANDOLFI PIER P. 2002. Modelling haematopoietic malignancies in the mouse and therapeutical implication. *Oncogene.* 21(21): 3445-3458.
- BODMER D, ELEVELD M, KATER-BAATS E, JANSSEN I, JANSSEN B, WETERMAN M, SCHOENMAKERS E, NICKERSON M, LINEHAN M, ZBAR B, VAN KESSEL AG. 2002. Disruption of a novel MFS transporter gene, DIRC2, by a familial renal cell carcinoma-associated t(2;3)(q35;q21). *Hum Mol Genet.* 15;11(6):641-9.
- BODMER D, JANSSEN I, JONKERS Y, VAN DEN BERG E, DIJKHUIZEN T, DEBIEC-RYCHTER M, SCHOENMAKERS E, VAN KESSEL AG. 2002. Molecular cytogenetic analysis of clustered sporadic and familial renal cell carcinoma-associated 3q13 approximately q22 breakpoints. *Cancer Genet Cytogenet.* 136(2):95-100.

- BORIES JC, CAYUELA JM, LOISEAU P, SIGAUX F. 1991. Expression of human recombination activating genes (RAG1 and RAG2) in neoplastic lymphoid cells: correlation with cell differentiation and antigen receptor expression. *Blood*. 78(8):2053-61.
- BROWN CK, MADAUSS K, LIAN W, BECK MR, TOLBERT WD, RODGERS DW. 2001. Structure of neurolysin reveals a deep channel that limits substrate access. *Proc Natl Acad Sci U S A*. 98(6):3127-32. Epub.
- CAMPESE AF, BELLAVIA D, GULINO A, SCREPANTI I. 2003. Notch signalling at the crossroads of T cell development and leukemogenesis. *Semin Cell Dev Biol*. 14(2):151-7.
- CARBONE A, GLOGHINI A, ZAGONEL V, ALDINUCCI D, GATTEI V, DEGAN M, IMPROTA S, SORIO R, MONFARDINI S, PINTO A. 1995. The expression of CD26 and CD40 ligand is mutually exclusive in human T-cell non-Hodgkin's lymphomas/leukemias. *Blood*. 86(12):4617-26.
- CARTLIDGE RA, KNEBEL A, PEGGIE M, ALEXANDROV A, PHIZICKY EM, COHEN P. 2005. The tRNA methylase METTL1 is phosphorylated and inactivated by PKB and RSK in vitro and in cells. *EMBO J*. 24(9):1696-705.
- CHABANNON C, MANNOUNI P. 1995. Les cellules souches hématopoïétiques du sang périphérique chez l'homme, *Médecine/science* 11(1):17-27.
- CHAO SH, CHENG TH, SHAW CY, LEE MH, HSU YH, TSAI YC. 2006. Characterization of a novel PepF-like oligopeptidase secreted by *Bacillus amyloliquefaciens* 23-7A. *Appl Environ Microbiol*. 72(1):968-71.
- CHOQUET S. *Hématologie*. 2002. Paris : ellipses, (Reussir l'internat).
- COFFIN JM, HUGBES SH, VARMUS HE. 1997. *Rétroviruses*. Cold Spring Harbor laboratory , NY. [Link to Retroviruses at the National Center for Biotechnology Information website].
- CONLEY YP, ERTURK D, KEVERLINE A, MAH TS, KERAVALA A, BARNES LR, BRUCHIS A, HESS JF, FITZGERALD PG, WEEKS DE, FERRELL RE, GORIN MB. 2000. Links A juvenile-onset, progressive cataract locus on chromosome 3q21-q22 is associated with a missense mutation in the beaded filament structural protein-2. *Am J Hum Genet*. 66(4):1426-31.
- CORCORAN LM, ADAMS JM, DUNN AR, CORY S. 1984. Murine T lymphomas in which the cellular myc oncogene has been activated by retroviral insertion. *Cell*. 37(1): 113-122.

- CURTIN JA, QUINT E, TSIPOURI V, ARKELL RM, CATTANACH B, COPP AJ, HENDERSON DJ, SPURR N, STANIER P, FISHER EM, NOLAN PM, STEEL KP, BROWN SD, GRAY IC, MURDOCH JN. 2003. Mutation of *Celsr1* disrupts planar polarity of inner ear hair cells and causes severe neural tube defects in the mouse. *Curr Biol*. 13(13):1129-33.
- DALLA-FAVERA R, BREGNI M, ERIKSON J, PATTERSON D, GALLO RC, CROCE CM. 1982. Human c-myc oncogene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc Natl Acad Sci USA*. 79(24): 7824-7.
- DENICOURT C, KOZAK CA, RASSART E. 2003. *Gris 1*, a new common Integration Site in Graffi Murine Leukemia Virus-Induced leukemia: Overexpression of a Truncated cyclin D2 due to Alternative Splicing. *J Virol*. 77(1):37-44.
- DENICOURT C, EDOUARD E, RASSART E. 1999. Oncogene Activation in Myeloïd leukemias by Graffi Murine Leukemia Virus Proviral Integration. *J Virol*. 73: 4439-4442.
- DEYHOLOS AND GALBRAITH. 2001; Geiss et al., 2001. La puce à ADN en médecine et en science : Le développement récent de la technique des microarrays a révolutionné l'étude de l'expression des gènes.
- DIOP S, LETESTU R, ORSOLANI D, LEOUEUF Y, LE TUTOUR P, THIAM D, DIAKHATE L, VALENSI F. 2005. Expression of proliferation marker Ki 67 in chronic lymphocytic leukemia. *Dakar Med*. 50(2):65-8.
- DITTEL BN, MCCARTHY JB, WAYNER EA, LEBIEN TW. 1993. Regulation of human B-cell precursor adhesion to bone marrow stromal cells by cytokines that exert opposing effects on the expression of vascular cell adhesion molecule-1. *Blood*. 81(9):2272-82.
- DONOHUE ME, BLOMBERG BB. 1997. The 14.1 surrogate light chain promoter has lineage- and stage-restricted activity. *J Immunol*. 158(4):1681-91.
- DUNN TB. 1954. Normal and pathologic anatomy of the reticular tissue in laboratory mice. *J. Natl Cancer. Inst*. 14(6):1281-1433.
- EISEN MB, SPELLMAN PT, BROWN PO, BOTSTEIN D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A*. 95(25):14863-8.
- FERRANDO AA, NEUBERG DS, STAUNTON J, LOH ML, HUARD C, RAIMONDI SC, BEHM FG, PUI CH, DOWNING JR, GILLILAND DG, LANDER ES, GOLUB TR, LOOK AT. 2002. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. *Cancer Cell*. 1(1):75-87.

- FEREIRA R, OHNEDA K, YAMAMOT OM, PHILIPSEN S. 2005. GATA-1 function, a paradigm for transcription factors in hematopoiesis. *Mol Cell Biol.* 25(4):1215-27.
- FEY, F, GRAFFI A. 1965. Erythroblastic leukemia after the injection of mouse myelogenous leukemia virus. *Z. Krebsforsch.* 67(2):145-151.
- FLANDRIN G. La nouvelle classification OMS des hémopathies malignes. www.biocolleges.org/colhemato/Hemato/Whogf.html.
- FREDRICKSON TN, HARRIS AW. 2000. Atlas of Mouse Hematopathology. Harwood academic publishers.
- FRIEDMAN AD. 2002. Transcriptional regulation of granulocyte and monocyte development. *Oncogene.* 21(21):3377-3390.
- GISSELBRECHT S. 2003. Oncogenes and leukemia: history and perspectives. *Med Sci (Paris).* 19(2):201-10.
- GRAFFI A, FEY F, AND T SCHRAMM. 1966. Experiments on the hematologic diversification of viral mouse leukemia. *Natl. Cancer Inst. Monogr.* 22:21-31.
- GRANGER SW, BUNDY LM, FAN H. 1999. Tandemization of a subregion of the enhancer sequences from SRS 19-6 murine leukemia virus associated with T-lymphoid but not other leukemias. *J Virol.* 73(9):7175-84.
- HANSEL DE, RAHMAN A, HOUSE M, ASHFAQ R, BERG K, YEO CJ, MAITRA A. 2004. Met proto-oncogene and insulin-like growth factor binding protein 3 overexpression correlates with metastatic ability in well-differentiated pancreatic endocrine neoplasms. *Clin Cancer Res.* 10(18 Pt 1):6152-8.
- HOFFMAN EC, REYES H, CHU FF, SANDER F, CONLEY LH, BROOKS BA, HANKINSON O. 1991. Cloning of a factor required for activity of the Ah (dioxin) receptor. *Science.* 252(5008):954-8.
- <http://anne.decoستر.free.fr/d1viro/vretrov0.html>
- <http://www.anne.decoستر.free.fr/d1viro/vtelechar/vpoly/hiv05.pdf>
- <http://www.chups.jussieu.fr/polys/cancero/POLY.Chp.3.5.html>
- HUNG MS, AVNER P, ROGNER UC. 2006. Identification of the transcription factor ARNTL2 as a candidate gene for the type 1 diabetes locus Idd6. *Hum Mol Genet.* 15(18):2732-42. Epub 2006.

- IKEDA M, YU W, HIRAI M, EBISAWA T, HONMA S, YOSHIMURA K, HONMA KI, NOMURA M. 2000. cDNA cloning of a novel bHLH-PAS transcription factor superfamily gene, BMAL2: its mRNA expression, subcellular distribution, and chromosomal localization. *Biochem Biophys Res Commun.* 275(2):493-502.
- INOUE M, SHEN GP, CHAUDHRY MA, GALICK H, BLAISDELL JO, WALLACE SS. 2004. Expression of the oxidative base excision repair enzymes is not induced in TK6 human lymphoblastoid cells after low doses of ionizing radiation. *Radiat Res.* 161(4):409-17.
- JAKOBSSON A, WESTERBERG R, JAKOBSSON A. 2006. Fatty acid elongases in mammals: their regulation and roles in metabolism. *Prog Lipid Res.* 45(3):237-49.
- JELINEK DF, TSCHUMPER RC, STOLOVITZKY GA, ITURRIA SJ, TU Y, LEPRE J, SHAH N, KAY NE. 2003. Identification of a global gene expression signature of B-chronic lymphocytic leukemia. *Mol Cancer Res.* 1(5):346-61.
- JOHNSTONE CN, CASTELLVI-BEL S, CHANG LM, BESSA X, NAKAGAWA H, HARADA H, SUNG RK, PIQUE JM, CASTELLS A, RUSTGI AK. 2004. ARHGAP8 is a novel member of the RHOGAP family related to ARHGAP1/CDC42GAP/p50RHOGAP: mutation and expression analyses in colorectal and breast cancers. *Gene.* 336(1):59-71.
- KATOH M. 2005. WNT/PCP signaling pathway and human cancer (review). *Oncol Rep.* 14(6):1583-8.
- KATOH Y, KATOH M. M&M MEDICAL BIOINFORMATICS. 2006. Comparative integromics on FAT1, FAT2, FAT3 and FAT4. *Int J Mol Med.* 18(3):523-8.
- KATZ SG, CANTOR AB, ORKIN SH. 2002. Interaction between FOG-1 and the corepressor C-terminal binding protein is dispensable for normal erythropoiesis in vivo. *Mol Cell Biol.* 22(9):3121-8.
- KLEIN U, TU Y, STOLOVITZKY GA., MATTIOLI M, CATTORETTI G, HUSSON H, FREEDMAN A, INGHIRAMI G, CRO L, BALDINI L, NERI A, CALIFANO A, DALLA-FAVERA R. 2001. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J. Exp. Med.* 194(11): 1625-1638.
- KOGAN SC, WARD JM, ANVER MR., BERMAN JJ, BRAYTON C, CARDIFF RD, CARTER JS, DE CORONADO S, DOWNING JR, FREDRICKSON TN, HAINES DC, HARRIS AW, HARRIS NL, HIAI H, JAFFE ES, MACLENNAN LAN CM, PANDOLFI PP, PATTENGAL PK, PERKINS AS, SIMPSON R M, TUTTLE MS, WONG JF, MRSEIII HC. 2002. Bethesda proposals for classification of non lymphoid hematopoietic neoplasms in mice. *Blood.* 100(1): 238-245.

- KORZ C, PSCHERER A, BENNER A, MERTENS D, SCHAFFNER C, LEUPOLT E, DOHNER H, STILGENBAUER S, LICHTER P. 2002. Evidence for distinct pathomechanisms in B-cell chronic lymphocytic leukemia and mantle cell lymphoma by quantitative expression analysis of cell cycle and apoptosis-associated genes. *Blood*. 99(12):4554-61.
- KOSARI F, ASMANN YW, CHEVILLE JC, VASMATZIS G. 2002. Cysteine-rich secretory protein-3: a potential biomarker for prostate cancer. *Cancer Epidemiol Biomarkers Prev*. 11(11):1419-26.
- LI L, WANG J, COOPER MD. 1993. cDNA cloning and expression of human glutamyl aminopeptidase (aminopeptidase A). *Genomics*. 17(3):657-64.
- LIU P, KELLER JR, ORTIZ M, TESSAROLLO L, RACHEL RA, NAKAMURA T, JENKINS NA, COPELAND NG. 2003. Bcl11a is essential for normal lymphoid development. *Nat Immunol*. 4(6):525-32. Epub 2003 Apr 28.
- LOOK AT. 1997. Oncogenic transcription factors in the human acute leukemias. *Science*. 278(5340):1059-64.
- LORD DUBE H, L'ITALIEN R. 1983. *Hématologie*. Décarie éditeur Montréal, Moline, éditeur Paris. Dépôt légal : 2^{ème} trimestre.
- MAECKER H, THOMAS JT, CONSORTI RL, ELLIS TM. 1996. A reversible defect in alpha-beta T cell receptor assembly. *Exp Cell Res*. 25;223(1):149-54.
- MAEDA T, LEM J, PALCZEWSKI K, HAESELEER F. 2005. A critical role of CaBP4 in the cone synapse. *Invest Ophthalmol Vis Sci*. 46(11):4320-7.
- MASMOUDI H, BEN AMMAR EA. 2000. *Immunologie fondamentale*. Centre de Publication Universitaire.
- MELHEM R, HAILAT N, KUICK R, HANASH SM. 1997. Quantitative analysis of Op18 phosphorylation in childhood acute leukemia. *Leukemia*. 11(10):1690-5.
- MORSE III HC, ANVER MR, FREDRICKSON TN, HAINES DC, HARIS AW, HARRIS NL, ELAINE SJ, KOGAN SC, MACLENNAN LCM, PATTENGAL PK, WARD JM. 2002. Bethesda proposals for classification of lymphoid neoplasms in mice. *Blood*. 100(1):246-258.
- NIEWIADOMSKA H, MIROWSKI M, KULCZYCKA D, NAJDER M, BALCERCZAK E, BLONSKI JZ. 2000. Some oncogene and tumour suppressor gene protein products expression in B-cell chronic lymphocytic leukaemia. *Cytobios*. 103(404):159-68.

- OHNISHI K, TAKEMORI T. 1994. Molecular components and assembly of mu.surrogate light chain complexes in pre-B cell lines. *J Biol Chem.* 269(45):28347-53.
- OKANO T, YAMAMOTO K, OKANO K, HIROTA T, KASAHARA T, SASAKI M, TAKANAKA Y, FUKADA Y. 2001. Chicken pineal clock genes: implication of BMAL2 as a bidirectional regulator in circadian clock oscillation. *Genes Cells.* 6(9):825-36.
- ONO Y, FUKUHARA N, YOSHIE O. 1997. Transcriptional activity of TAL1 in T cell acute lymphoblastic leukemia (T-ALL) requires RBTN1 or -2 and induces TALLA1, a highly specific tumor marker of T-ALL. *J Biol Chem.* 272(7):4576-81.
- PALMER MT, MORROW CD. 2004. Analysis of murine leukemia virus replication complemented by yeast tRNA^{phc} reveals in herent preferences for the tRNA primer selected for reverse transcriptase. *J Virol.* 324(2):430-438.
- CORDELIER D. 1975. *Immunologie fondamentale*. Tome1. Edition Crouan et Roques.
- PARK JS, ONODERA T, NISHIMURA S, THOMPSON RF, ITOHARA S. 2006. Molecular evidence for two-stage learning and partial laterality in eyeblink conditioning of mice. *Proc Natl Acad Sci U S A.* 103(14):5549-54.
- RASSART E. 2004:
http://www.unites.uqam.ca/biologie_moleculaire/images/retro.html
- REDDY P, JACQUIER AC, ABOVICH N, PETERSEN G, ROSBASH M. 1986. The period clock locus of *D. melanogaster* codes for a proteoglycan. *Cell.* 46(1):53-61.
- ROWEN L, YOUNG J, BIRDITT B, KAUR A, MADAN A, PHILIPPS DL, QIN S, MINX P, WILSON RK, HOOD L, GRAVELEY BR. 2000. Analysis of the human neurexin genes: alternative splicing and the generation of protein diversity. *Genomics.* 79(4):587-97.
- RU M., SHUSTIK C., AND RASSART ERIC. 1993. Graffi Murine Leukemia Virus: Molecular Cloning and Characterization of the Myeloid Leukemia-Inducing Agent. *J Virol.* 67(8): 4722-4731.
- RYLEY DA, WU HH, LEADER B, ZIMON A, REINDOLLAR RH, GRAY MR. 2005. Characterization and mutation analysis of the human formin-2 (FMN2) gene in women with unexplained infertility. *Fertil Steril.* 83(5):1363-71.
- SAWYERS CL, DENNY CT, WITTE ON. 1991. Leukemia and the Disruption of Normal Hematopoiesis. *Cell* 64(2):337-350.
- SELTEN G, CUYPERS HT, ZIJLSTRA M, MELIEF C, BERNIS A. 1984. Involvement of c-myc in MuLV-induced T cell lymphomas in mice: Frequency and mechanisms of activation. *EMBO J.* 3(13): 3215-3222.

- SHI CS, TUSCANO JM, WITTE ON, KEHRL JH. 1999. GCKR links the Bcr-Abl oncogene and Ras to the stress-activated protein kinase pathway. *Blood*. 93(4):1338-45.
- SHIMAMOTO T, OHYASHIKI K, OHYASHIKI JH, KAWAKUBO K, FUJIMURA T, IWAMA H, NAKAZAWA S, TOYAMA K. 1995. The expression pattern of erythrocyte/megakaryocyte-related transcription factors GATA-1 and the stem cell leukemia gene correlates with hematopoietic differentiation and is associated with outcome of acute myeloid leukemia. *Blood*. 86 (8):3173-80.
- SITBON M, DENESBRE C, DARDALHON V, BATTINI JL, MOUGEL M. 2001. Les retrovirus leucémogènes murins : pathogènes, gènes et outils génétiques. *Virol*. 5(4):265-293.
- STRASSBURG C.P, NEUBAUER V, POLIWOADA H, BENTER T. 1992. Regulation of the proto-oncogenes c-cis, c-fos, c-myc and c-myb in acute myeloid leukemia. *Neoplasma*. 39(6):343-347.
- STRATOWA C, LOFFLER G, LICHTER P, STILGENBAUER S, HABERL P, SCHWEIFER N, DOHNER H, WILGENBUS KK. 2001. CDNA microarray gene expression analysis of B-cell chronic lymphocytic leukemia proposes potential new prognostic markers involved in lymphocyte trafficking. *Int J Cancer*. 91(4):474-80.
- SUN Y, DONG LJ, TIAN F, WANG SQ, JIA ZL, HUANG J, CHEN ZJ, LI WJ, CHEN XL, ZHU P. 2004. Identification of acute leukemia-specific genes from leukemia recipient/sibling donor pairs by distinguishing study with oligonucleotide microarrays. *Zhongguo Shi Yan Xue Ye Xue Za Zhi*. 12(4):450-4.
- TAUB R, KIRSCH I, MORTON C, LENOIR G, SWAN D, TRONICK S, AARONSON S, LEDER P. 1982. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cell. *Proc Natl Acad Sci USA*. 79(24):7837-41.
- TAYLOR BL, ZHULIN IB, JOHNSON MS. 1999. Aerotaxis and other energy-sensing behavior in bacteria. *Annu Rev Microbiol*. 53:103-28.
- TORISU K, TSUCHIMOTO D, OHNISHI Y, NAKABEPPU Y. 2005. Hematopoietic Tissue-Specific Expression of Mouse Neil3 for Endonuclease VIII-Like Protein. *J Biochem (Tokyo)*. 138(6):763-72.
- TRENTIN L, CABRELLE A, FACCO M, CAROLLO D, MIORIN M, TOSONI A, PIZZO P, BINOTTO G, NICOLARDI L, ZAMBELLO R, ADAMI F, AGOSTINI C, SEMENZATO G. 2004. Homeostatic chemokines drive migration of malignant B cells in patients with non-Hodgkin lymphomas. *Blood*. 104(2):502-8. Epub 2004 Mar 4.

- VOISIN V, BARAT C, HOANG T, RASSART E. 2006. Novel Insights into the Pathogenesis of the Graffi Murine Leukemia Retrovirus. *J virol.* 80(8):4026-4037.
- WECHSLER J, GREENE M, MCDEVITT MA, ANASTASI J, KARP JE, LE BEAU MM, CRISPINO JD. 2002. Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. *Nat Genet.* 32(1):148-52.
- WHITE LJ, OZANNE BW, GRABER P, AUBRY JP, BONNEFOY JY, CUSHLEY W. 1997. Inhibition of apoptosis in a human pre-B-cell line by CD23 is mediated via a novel receptor. *Blood.* 90(1):234-43.
- WIKMAN H, NYMARK P, VAYRYNEN A, JARMALAITIE S, KALLIONIEMI A, SALMENKIVI K, VAINIO-SIUKOLA K, HUSGAFVEL-PURSIAINEN K, KNUUTILA S, WOLF M, ANTTILA S. 2005. CDK4 is a probable target gene in a novel amplicon at 12q13.3-q14.1 in lung cancer. *Genes Chromosomes Cancer.* 42(2):193-9.
- World Health organization. Classification of Neoplasms Diseases of the Hematopoietic and Lymphoid Tissues CDRAFT. 1997. *American Journal of Surgical Pathology.* 21(1): 114-121.
www.dartmouth.edu/~nlevuy/WHO_classification.html
- YEH CT, LU SC, TSENG IC, LAI HY, TSAO ML, HUANG SF, LIAW YF. 2003. Antisense overexpression of BMAL2 enhances cell proliferation. *Oncogene.* 22(34):5306-14.
- YOSHIE O. 2005. Expression of CCR4 in adult T-cell leukemia. *Leuk Lymphoma.* 46(2):185-90.